Covariant differentiation of a map in the context of geometric optimal control

Alessandro Saccon * **John Hauser** ** **A. Pedro Aguiar** ***

* *Department of Mechanical Engineering, Eindhoven University of Technology, the Netherlands (e-mail:* `a.saccon@tue.nl`*)*
** *Department of Electrical, Computer, and Energy Engineering, University of Colorado Boulder (e-mail:* `hauser@colorado.edu`*)*
*** *Faculty of Engineering, University of Porto (FEUP) (e-mail:* `pedro.aguiar@fe.up.pt`*)*

Abstract:
This paper provides a detailed discussion of the second covariant derivative of a map and its role in the Lie group projection operator approach (a direct method for solving continuous time optimal control problems). We begin by briefly describing the iterative geometric optimal control algorithm and summarize the general expressions involved. Particular emphasis is placed on the expressions related to the search direction subproblem, writing them in a new compact form by using a new operator notation. Next, we show that the covariant derivative of a map between manifolds endowed with affine connections plays a key role in obtaining the required local quadratic approximations for the Lie group projection operator approach. We present a new result for computing an approximation of the parallel displacement associated with an affine connection which is an affine combination of two (or more) connections. As a corollary, an extremely useful approximation of the parallel displacement relative to the Cartan-Schouten (0) connection on Lie groups is obtained.

Keywords: Optimal control, Riccati equations, differential geometry, geometric approaches, Lie groups, Projection operator approach

1. INTRODUCTION

In Saccon et al. (2013), we have proposed an algorithm for solving continuous time optimal control problems for systems evolving on (noncompact) Lie groups (including, as a particular case, the flat space $\mathbb{R}^n$). The approach borrows from and expands the key results of the projection operator approach for the optimization of trajectory functionals, developed in Hauser (2002).

The algorithm can be viewed as a generalization of Newton's method to the infinite dimensional setting and exhibits a second order convergence rate to a local minimizer at which the second order sufficient condition for optimality holds. At each step, a quadratic model of a cost functional (given by the composition of the original cost functional with the projection operator) is constructed about the current trajectory iterate. This quadratic model is developed using the first and second derivatives of the incremental cost, terminal cost, and the control system vector field. To this end, the second covariant derivative of a map between two manifolds plays a key role in providing a chain rule for the required Lie group computations.

Motivated by this, we provide technical details and a historical perspective of the second covariant derivative of a map between smooth manifolds endowed with affine connections. We present a new result for computing an approximation of the parallel displacement associated with an affine connection which is an affine combination of two (or more) connections. As a corollary, an extremely useful approximation of the parallel displacement relative to the Cartan-Schouten (0) connection on Lie groups is obtained. Having at hand such an approximation of the parallel displacement is key for computing the second covariant derivatives of the cost function and system dynamics that are used in the projection operator approach.

The paper is organized as follows. Section 2 introduces the notation used throughout the paper. The projection operator approach on Lie groups for the optimization of trajectory functionals is reviewed in Section 3 and rewritten in a new compact form by using operator theory. In Section 4, historical development and important properties of the second covariant derivative of a map are detailed. Finally, in Section 5 an approximation of the parallel displacement relative to the (0) Cartan-Schouten connection on Lie groups is introduced and justified. Concluding remarks are given in Section 6.

⋆ This work was supported in part by projects CONAV/FCT-PT (PTDC/EEACRO/113820/2009), FCT (PEst-OE/EEI/LA0009/2011), and MORPH (EU FP7 under grant agreement No. 288704). The work of the second author was supported in part by AFOSR FA9550-09-1-0470 and by an invited scientist grant from the Foundation for Science and Technology (FCT), Portugal. The first author benefited from a postdoctoral fellowship from FCT.

2. NOTATION

This section has been reduced in size in the final submission due to page limitation. Please contact the authors to receive a complete version of it.

Notation

M, N	Smooth manifolds
T_xM, T_x^*M	Tangent and cotangent spaces of M at $x \in M$
TM, T^*M,	Tangent and cotangent bundles of M
$\phi : M \to N$	A map (a function, when $N = \mathbb{R}$)
$\mathbf{D}\,\phi{:}TM \to TN$	Tangent map of ϕ (the differential of ϕ, when $N = \mathbb{R}$)
$\mathfrak{L}$	set of linear maps between two vector spaces, e.g., $\mathfrak{L}(T_xM, T_{\phi(x)}N)$
∇	Affine connection on a smooth manifold
$\nabla_X Y$	Covariant derivative of a vector field Y in the direction X
$\gamma : \mathbb{R} \to M$	A smooth curve on M
$P_\gamma^{t_1 \leftarrow t_0} V_0$	Parallel displacement along γ of the vector $V_0 \in T_{\gamma(t_0)}M$ from $\gamma(t_0)$ to $\gamma(t_1)$
$\mathbb{D}^2\phi(x) \cdot (\mathrm{v}_1, \mathrm{v}_2)$	Second covariant derivative of ϕ at x, evaluated in the directions v_1, $\mathrm{v}_2 \in T_xM$
$\operatorname{Hess}\phi(x)$	Hessian of ϕ at x
G	a Lie group
$\mathfrak{g}$	the Lie algebra of G
e	Group identity
L_g, R_g	Left and right translations by $g \in G$
TL_g, TR_g	Tangent maps of L_g and R_g, respectively (evaluated at e)
$[\cdot, \cdot]$	Lie bracket operation
Ad	Adjoint representation of G on $\mathfrak{g}$
ad	Adjoint representation of $\mathfrak{g}$ onto itself, satisfying $\mathrm{ad}(\varrho)(\varsigma) = [\varrho, \varsigma]$
$\mathrm{ad}^{\leftrightharpoons}$	ad operator with swapped arguments, $\mathrm{ad}^{\leftrightharpoons}(\varrho)(\varsigma) = \mathrm{ad}(\varsigma)(\varrho)$
$\exp : \mathfrak{g} \to G$	Exponential map
$\log : G \to \mathfrak{g}$	Logarithm map (inverse of the exp in a neighborhood of e)

The second covariant derivative of ϕ at $x_0 \in M$ evaluated in the directions v_1, $\mathrm{v}_2 \in T_{x_0}M$ is computed as

$$\mathbb{D}^2\phi(x_0) \cdot (\mathrm{v}_1, \mathrm{v}_2) = \lim_{t_1 \to t_0} \frac{1}{t_1 - t_0}\Big({}^N P_{\phi\circ\gamma}^{t_0 \leftarrow t_1} \mathbf{D}\,\phi(x_1) \cdot {}^M P_\gamma^{t_1 \leftarrow t_0} \mathrm{v}_1 - \mathbf{D}\,\phi(x_0) \cdot \mathrm{v}_1\Big). \tag{1}$$

3. THE PROJECTION OPERATOR APPROACH

The projection operator approach, detailed in Hauser (2002), is a direct method for solving continuous time optimal control problems generating a sequence of trajectories with decreasing cost. It differs from many standard numerical methods for solving optimal control problems in that it does not make use of a transcription phase wherein the system dynamics and cost functional are discretized to obtain a (finite dimensional) nonlinear optimization problem. Rather, all computations are accomplished in the space of continuous time trajectories and curves. Each iteration of the algorithm amounts to integrating the system dynamics and solving an associated set of linear and Riccati equations by means of an ordinary differential equation solver.

3.1 The method in a nutshell

In Saccon et al. (2010) and Saccon et al. (2013), the authors have shown how the trajecotry optimization approach presented in Hauser (2002) can be generalized to work with a dynamical system that evolves on a Lie group G, that is, for a system of the form

$$\dot{g} = f(g, u, t) = g(t)\lambda(g(t), u(t), t)\,, \tag{2}$$

where $f : G \times \mathbb{R}^m \times \mathbb{R} \to TG$ is a control vector field on G and $\lambda : G \times \mathbb{R}^m \times \mathbb{R} \to \mathfrak{g}$, $\lambda(g, u, t) := g^{-1} f(g, u, t)$, is its *left-trivialization*. Implementations details and numerical experiments on SO(3), TSO(3), and SE(3) have been reported, respectively, in Saccon et al. (2011b) Saccon et al. (2011a) and Saccon et al. (2012).

The approach, in its simplest formulation, can handle optimal control problems of the form

$$\min_{(g,u)(\cdot)} \int_0^{t_f} l(g(\tau), u(\tau), \tau)\, d\tau + m(g(t_f)) \tag{3}$$

subject to

$$\dot{g} = f(g, u, t), \tag{4}$$

$$g(0) = g_0, \tag{5}$$

where $l : G \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$ is the incremental cost, $m : G \to \mathbb{R}$ the terminal cost, and g_0 the initial condition. Modifications of the strategy for handling a terminal condition and mixed input-state constraints (through a barrier function approach) are discussed, for control problems on $\mathbb{R}^n$, in Hauser (2003) and Hauser and Saccon (2006). A constrained optimal control problem on Lie groups is discussed in Saccon et al. (2012).

Roughly speaking, the projection operator approach can be thought as a Newton descent method in infinite dimensions. The method is based on (and derives its name from) the projection operator $\mathcal{P}$, an operator that maps a generic state-input *curve* $\xi(t) = (\alpha(t), \mu(t)) \in G \times \mathbb{R}^m$, $t > 0$, into a state-input *trajectory* $\eta(t) = (g(t), u(t)) \in G \times \mathbb{R}^m$, $t > 0$, of the system (2). The set of system trajectories $\mathcal{T}$ is an infinite dimensional Banach manifold (Hauser and Meyer, 1998).

On a Lie group, the operator $\mathcal{P}$ is defined as

$$\begin{aligned} \dot{g}(t) &= g(t)\lambda(g(t), u(t), t)\,, \quad g(0) = \alpha(0)\,, \\ u(t) &= \mu(t) + K_r(t)\big[\log(g(t)^{-1}\alpha(t))\big]\,, \end{aligned} \tag{6}$$

where the regulator gain $K_r(t) : \mathfrak{g} \to \mathbb{R}^m$ is a linear map, which can be thought as a standard linear feedback once a basis is chosen for the Lie algebra $\mathfrak{g}$. It is straightforward to verify that $\mathcal{P}$ is indeed a projection, satisfying $\mathcal{P}^2 = \mathcal{P} \circ \mathcal{P} = \mathcal{P}$.

Given a trajectory $\xi(\cdot) = (g(\cdot), u(\cdot))$ of the control system (2), we define its (left-trivialized) linearization to be the time-varying linear system

$$\dot{z}(t) = A(\xi(t), t)z(t) + B(\xi(t), t)v(t), \tag{7}$$

with $(z(t), v(t)) \in \mathfrak{g} \times \mathbb{R}^m$, $t \geq 0$ and where

$$A(\xi, t) := A((g, u), t) = \mathbf{D}_1\, \lambda(g, u, t) \circ T_e L_g - \mathrm{ad}_{\lambda(g,u,t)}, \tag{8}$$

$$B(\xi, t) := B((g, u), t) = \mathbf{D}_2\, \lambda(g, u, t). \tag{9}$$

Let h be the cost functional appearing in (3) and define $\tilde{h}$ to be the functional obtained by composing h with the projection operator $\mathcal{P}$, i.e., $\tilde{h} := h \circ \mathcal{P}$. In this fashion, $\tilde{h}$ is a trajectory functional that incorporates the dynamic trajectory constraint (4)-(5) into the cost in an *unconstrained* manner.

The projection operator approach consists in applying the following iterative method

Algorithm (Projection operator Newton method)
given initial trajectory $\xi_0 \in \mathcal{T}$
for $i = 0, 1, 2, \ldots$

(search direction subproblem)

$$\zeta_i = \arg \min_{\xi_i \zeta \in T_{\xi_i} \mathcal{T}} \mathbf{D}\, h(\xi_i) \cdot \xi_i \zeta + \tfrac{1}{2} \mathbb{D}^2 \tilde{h}(\xi_i) \cdot (\xi_i \zeta, \xi_i \zeta) \tag{10}$$

$$\gamma_i = \arg \min_{\gamma \in (0,1]} \tilde{h}(\xi_i \exp(\gamma \zeta_i)) \qquad \text{(step size)} \tag{11}$$

$$\xi_{i+1} = \mathcal{P}(\xi_i \exp(\gamma_i \zeta_i)) \qquad \text{(update)} \tag{12}$$

end

In each iteration, the search direction minimization (10) is performed on the tangent space $T_\xi \mathcal{T}$, that is, we search over curves $\zeta(\cdot) = (z(\cdot), v(\cdot))$ satisfying (7). Then, the step size subproblem (11) is considered. The classical *approximate* solution obtained using a backtracking line search with Armijo condition (see, e.g., Chapter 3 of Nocedal and Wright (2006)) can be used to compute an appropriate step size γ_i. Finally, the update computation (12) *projects* the iterate onto the trajectory manifold and the iteration proceeds until a termination condition has been met.

In Saccon et al. (2011b) and Saccon et al. (2013), we have shown that the search direction step of the projection operator amounts to solving an associated linear-quadratic (LQ) regulator problem. In the following, we provide, for the first time, all of the details of the ordinary differential equations that need to be solved for obtaining the solution to this particular LQ problem. Such a detailed description should facilitate the comprehension of the practical aspects of the method and provide a useful reference and, indeed, a *cheat sheet* for its implementation.

3.2 Search direction subproblem

Given a trajectory $\xi = (g(\cdot), u(\cdot)) \in \mathcal{T}$, the search direction subproblem amounts to finding $\zeta = (z(\cdot), v(\cdot))$ that satisfies

$$\zeta = \arg \min_{\xi\zeta \in T_\xi \mathcal{T}} \mathbf{D}\, h(\xi) \cdot \xi\zeta + (1/2)\, \mathbb{D}^2 \tilde{h}(\xi) \cdot (\xi\zeta, \xi\zeta).$$

As shown in Saccon et al. (2011b) and Saccon et al. (2013), the search direction step (10) is equivalent to solving an optimal control problem of the form

$$\min_{(z,v)(\cdot)} \int_0^{t_f} a(\tau)^T z(\tau) + b(\tau)^T v(\tau) + \frac{1}{2} \begin{bmatrix} z(\tau) \\ v(\tau) \end{bmatrix}^T W(\tau) \begin{bmatrix} z(\tau) \\ v(\tau) \end{bmatrix} d\tau + r_1^T z(t_f) + \frac{1}{2} z(t_f)^T P_1 z(t_f)\,, \tag{13}$$

subject to the dynamic constraint

$$\dot{z} = A(\xi(t), t) z + B(\xi(t), t) v, \qquad z(0) = 0, \tag{14}$$

with $z(t) \in \mathfrak{g}$ and $v(t) \in \mathbb{R}^m$. The expressions for the vectors $a(t)$, $b(t)$, r_1, and the matrices $A(\xi(t), t)$, $B(\xi(t), t)$, $W(t)$ and P_1 above, are detailed below.

The linear quadratic optimal control problem (13)-(14) can be solved by stardard techniques (Anderson and Moore, 1989), (Bryson and Ho, 1969). The linear two-point boundary value problem obtained from the first order necessary condition (alternately, the Euler-Lagrange equations or Pontryagin Maximum Principle) can be, when solvable, converted to a set of ordinary differential equations running backward and forward in time by using a Riccati transformation (cf. the *sweep* method in Bryson and Ho (1969)).

Integration backward in time. A standard Riccati differential equation associated with the LQ optimal control problem is solved backward in time. Furthermore, an extra linear differential equation (see (17) below) is needed, due to the presence of the linear term $\int_0^{t_f} a^T z + b^T v \, d\tau + r_1^T z(t_f)$ in the cost functional (13). In detail, we solve

$$K_o = R_o^{-1}(S_o^T + B^T P), \tag{15}$$

$$\dot{P} = A^T P + PA - K_o^T R_o K_o + Q_o, \quad P(t_f) = P_1, \tag{16}$$

$$-\dot{r} = (A - BK_o)^T r + a - K_o^T b, \qquad r(t_f) = r_1, \tag{17}$$

$$-\dot{q} = (A - BK_r)^T q + a - K_r^T b, \qquad q(t_f) = r_1, \tag{18}$$

$$v_o = -R_o^{-1}(B^T r + b), \tag{19}$$

with

$$a = l_g^T, \qquad b = l_u^T, \tag{20}$$

$$Q_o = l_{gg} + \sum q_k \lambda_{k,gg}, \quad S_o = l_{gu} + \sum q_k \lambda_{k,gu}, \tag{21}$$

$$R_o = l_{uu} + \sum q_k \lambda_{k,uu}, \tag{22}$$

$$r_1 = m_g^T, \qquad P_1 = m_{gg}. \tag{23}$$

The expressions above are time-varying in general and are evaluated along the trajectory $\xi(t) = (g(t), u(t)) \in G \times \mathbb{R}^m$, $t \in [0, t_f]$. The exceptions are the time invariant expressions r_1 and P_1 which are evaluated at $g(t_f)$.

The state q, appearing in (18) and component by component in (21)-(22), is a stabilized 'adjoint' state associated with the second covariant derivative of the projection operator, detailed in Saccon et al. (2011b) and Saccon et al. (2013), and in Hauser (2002) for the flat case. Interestingly, this Lagrange multiplier arises *automagically* in the computation of $\mathbb{D}^2 \tilde{h}(\xi_i) \cdot (\xi_i \zeta, \xi_i \zeta)$ *without* reference to any stationarity conditions! In (15), K_o denotes the optimal gain associated with the solution of (13)-(14), while the gain K_r in (18) is that associated with the projection operator $\mathcal{P}$ in (6).

When the method is applied to a dynamical system whose state evolves on $\mathbb{R}^n$, the terms $l_g, l_u, l_{gg}, l_{gu}, l_{gg}, \ldots$, found in (20)-(23) refer to standard first and second (mixed) partial derivatives with respect to the state and input variables. On general Lie groups, these expressions are slightly more complicated due to the presence of extra terms involving left trivialization and the geometry of the state manifold.

The general expressions to be used in (15)-(23), involving the first and second derivatives of the incremental cost l, final cost m, and the dynamics λ, are the following:

$$A = \mathbf{D}_1\,\lambda \circ TL_g - \mathrm{ad}_\lambda, \tag{24}$$
$$B = \mathbf{D}_2\,\lambda \tag{25}$$

$$\begin{aligned}\lambda_{k,gg} &= (TL_g)^* \circ \mathrm{Hess}_1\,\lambda_k \circ TL_g \\ &\quad + 1/2\left(\mathrm{ad}^{\leftrightharpoons} \circ \mathbf{D}_1\,\lambda \circ TL_g\right)_k^{\mathfrak{g}} \\ &\quad + 1/2\left((TL_g)^* \circ (\mathbf{D}_1\,\lambda)^* \circ \mathrm{ad}\right)_k^{\mathfrak{g}} \end{aligned} \tag{26}$$
$$\lambda_{k,gu} = \left(\mathbf{D}_2(\mathbf{D}_1\,\lambda \circ TL_g) + 1/2\left(\mathrm{ad}^{\leftrightharpoons} \circ \mathbf{D}_2\,\lambda\right)\right)_k^{\mathfrak{g}} \tag{27}$$
$$\lambda_{k,ug} = \left(\mathbf{D}_1(\mathbf{D}_2\,\lambda) \circ TL_g + 1/2\left((\mathbf{D}_2\,\lambda)^* \circ \mathrm{ad}\right)\right)_k^{\mathbb{R}^m} \tag{28}$$
$$\lambda_{k,uu} = \mathrm{Hess}_2\,\lambda_k \tag{29}$$

$$l_g = \mathbf{D}_1\,l \circ TL_g \tag{30}$$
$$l_u = \mathbf{D}_2\,l \tag{31}$$
$$l_{gg} = (TL_g)^* \circ \mathrm{Hess}_1\,l \circ TL_g \tag{32}$$
$$l_{gu} = \mathbf{D}_2(\mathbf{D}_1\,l \circ TL_g) \tag{33}$$
$$l_{ug} = \mathbf{D}_1(\mathbf{D}_2\,l) \circ TL_g \tag{34}$$
$$l_{uu} = \mathrm{Hess}_2\,l \tag{35}$$

$$m_g = \mathbf{D}\,m \tag{36}$$
$$m_{gg} = (TL_g)^* \circ \mathrm{Hess}\,m \circ TL_g \tag{37}$$

In (26)-(29), $(\cdot)_k^E : \mathfrak{L}(E,\mathfrak{g}) \to \mathfrak{L}(E,\mathbb{R})$, $k \in \{1,2,\ldots,n\}$, $E = \{\mathfrak{g}, \mathbb{R}^m\}$, are defined as $(\eta)_k^E : \eta \mapsto \pi_k \circ \eta$, where $\pi_k : \mathfrak{g} \to \mathbb{R}$, $k \in \{1,2,\ldots,n\}$, are the natural projections a Lie algebra element into its components for the given choice of a basis for $\mathfrak{g}$.

Equations (24)-(37) were introduced in Saccon et al. (2011b) and further explained in Saccon et al. (2013). Here, they are presented all together in a new compact form obtained using the operator notation, as explained in the notation section. Recall that $\mathbb{R}^n$ with standard addition $+$ is a Lie group. Indeed, for a system evolving on $\mathbb{R}^n$, (24)-(37) simplify to the standard first and second derivatives of the dynamics and cost functions.

After the optimal gain K_o in (15) and the optimal feedforward v_o in (19) have been computed backward in time, the following forward time integration allows one to obtain the optimal state-input trajectory $\zeta(\cdot) = (z(\cdot), v(\cdot))$ that solves the LQ optimal control problem (13)-(14), together with its associated optimal value.

Integration Forward in time. Solve

$$\dot z = Az + Bv, \qquad z(0) = 0, \tag{38}$$
$$v = -K_o z + v_o, \tag{39}$$
$$\dot z_{n+1} = a^T z + b^T v, \qquad z_{n+1}(0) = 0, \tag{40}$$
$$\dot z_{n+2} = z^T Q_o z + 2 z^T S_o v + v^T R_o v, \qquad z_{n+2}(0) = 0. \tag{41}$$

The state z_{n+1} and z_{n+2} evaluated at t_f allow to compute

$$Dh(\xi) \cdot \xi\zeta = z_{n+1}(t_f) + r_1^T z(t_f) \tag{42}$$
$$\mathbb{D}^2 \tilde h(\xi) \cdot (\xi\zeta, \xi\zeta) = z_{n+2}(t_f) + z(t_f)^T P_1 z(t_f). \tag{43}$$

With the descent direction ζ_i in hand, local trajectories of the form $\xi = \mathcal{P}(\xi_i \exp(\gamma\zeta_i))$ are used in a backtracking "line" search on the trajectory manifold to obtain a suitable descent increment.

4. COVARIANT DERIVATIVES OF A MAP

In Saccon et al. (2010), we introduced the concept of the second geometric derivative of a map between manifolds endowed each with an affine connection. This mathematical object is a key ingredient in the projection operator approach as it is used in obtaining a local quadratic approximation of a given optimal control problem.

We later discovered and reported in Saccon et al. (2013) that this type of derivation of a map between differentiable manifolds is a special case of covariant differentiation of two-point tensor fields. Two-point tensor fields (sometimes also called double tensor fields) are not commonly encountered in standard differential and Riemannian geometry textbooks. They can be found in advanced continuum mechanics, quantum physics, and differential geometry applications and are natural generalization of vector fields and one forms over maps. An introduction to double tensor fields can be found in Ericksen (1960) and Marsden and Hughes (1983).

In Saccon et al. (2010), we also pointed out that the second covariant derivative of a map between two manifolds seemed to somehow be related to the classical concept of second fundamental form associated with an isometric embedding in the context of Riemannian geometry. We have recently discovered that in fact this intuition is correct. The second covariant derivative is indeed equivalent to the classical concept of second fundamental form of Riemannian geometry and can be extended to arbitrary maps between Riemannian manifolds, where it is usually referred to as *second fundamental form of a map*. According to Vilms (1970), the definition of the second fundamental form of a map is due to Eells and Sampson (1964), where it was used to generalize the concept of harmonic function by introducing the concept of harmonic map.

The key properties of the second fundamental form of a map have been detailed in Vilms (1970) and Eells and Lemaire (1978). For example, in Vilms (1970), a map with vanishing second covariant derivative is shown to be totally geodesic, characterized by the property that it carries geodesics to geodesics.

Some of the properties discussed in Vilms (1970) and Eells and Lemaire (1978) do not rely on the existence of a metric (on the domain and codomain manifolds) and therefore hold in general for the second covariant derivative of a map. When a set of local coordinates is chosen, one can use the explicit expression of the second covariant derivative of a map in terms of the Christoffel symbols associated to the connections as given in Section 3 of Eells and Lemaire (1978). Let $\phi : M \to N$ be a twice differentiable mapping between smooth manifolds M and N endowed with affine connections with Christoffel symbols denoted by ${}^M\Gamma^k_{ij}$, i, j, $k \in \{1,2,\ldots,\dim M\}$ and ${}^N\Gamma^\gamma_{\alpha\beta}$, α, β, $\gamma \in \{1,2,\ldots,\dim N\}$. Then, the second covariant derivative of the map ϕ is given component-wise by

$$\begin{aligned}&(\mathbb{D}^2\phi \cdot (e_i, e_j))^\gamma \\ &\qquad = \frac{\partial^2 \phi^\gamma}{\partial x^i \partial x^j} - {}^M\Gamma^k_{ij}\frac{\partial \phi^\gamma}{\partial x^k} + {}^N\Gamma^\gamma_{\alpha\beta}\frac{\partial\phi^\alpha}{\partial x^i}\frac{\partial\phi^\beta}{\partial x^j}.\end{aligned} \tag{44}$$

In Section 1 of Vilms (1970), it is further proven that the skew-symmetric part of the second covariant derivative of a map depends only on the torsion of the two connections. This latter results allows one to conclude that if both manifolds are endowed with symmetric connections, then the second covariant derivative is symmetric. This result is contained in Vilms (1970), of which we were not aware at the time we wrote Saccon et al. (2010), justifies the use of the torsion-free Cartan-Schouten (0) connection in Saccon et al. (2010) to obtain symmetric expressions for the local quadratic approximation of the optimal control problem.

5. AFFINE COMBINATION OF CONNECTIONS

Recall that the expression for computing the second covariant derivative of a map (1) makes explicit use of parallel displacement. When using the Cartan-Schouten (0) connection in Saccon et al. (2010), we erroneously indicated that its parallel displacement was path independent and given by

$$P_\gamma^{t_1 \leftarrow t_0} \mathrm{v}_0 = \frac{1}{2}\big(x_1 x_0^{-1} \mathrm{v}_0 + \mathrm{v}_0 x_0^{-1} x_1\big), \qquad (45)$$

with $\gamma : \mathbb{R} \to G$ a curve on the Lie group G, $x_0 = \gamma(t_0)$, $x_1 = \gamma(t_1)$, and $\mathrm{v}_0 \in T_{x_0}G$. The use of (45) in computing the second covariant derivative of a map relative to the (0) Cartan-Schouten connection is however justifiable (and also extremely useful!) since (45) is in fact a first order approximation of the *path dependent* parallel displacement of the (0) connection. Examining (1), it is clear that a first order (in time) approximation of the parallel transport is sufficient for computing the second covariant derivative.

Hereafter, we formally justify why (45) provides a valid first order approximate expression. We show that this follows from a more general result involving an affine combination of affine connections. The approximate expression (45) was presented as an approximation already in Saccon et al. (2013), although without the detailed and explicit justification that we now provide.

An affine combination is a linear combination where the coefficients sum to one. While a linear combination of connections is not necessarily a connection, an affine combination always is (Lee, 1997). The key is that the required product rule is not satisfied when the combination is not affine. The following proposition provides a tool for approximating the parallel displacement for a connection that is an affine combination of connections in terms of the parallel displacements of the constituent connections. To the best of our knowledge, this result is new.

Proposition 5.1. Let ${}^1\nabla$ and ${}^2\nabla$ be affine connections on a smooth manifold M, with corresponding parallel displacements ${}^1P$ and ${}^2P$, and define

$$\nabla := \alpha_1\, {}^1\nabla + \alpha_2\, {}^2\nabla \qquad (46)$$

with $\alpha_1 + \alpha_2 = 1$. Then, ∇ is a connection on M and the corresponding parallel displacement P satisfies

$$P_\gamma^{t_1 \leftarrow t_0} V_0 = \big(\alpha_1\, {}^1P_\gamma^{t_1 \leftarrow t_0} + \alpha_2\, {}^2P_\gamma^{t_1 \leftarrow t_0}\big) V_0 + o(t_1 - t_0), \qquad (47)$$

for any $V_0 \in T_{\gamma(t_0)}M$ with γ a smooth curve on M. (Here $o(\cdot)$ is the Landau symbol indicating that the remainder is 'little-oh' in $t_1 - t_0$.)

Proof: Given a set of local coordinates x^i, $i \in \{1, 2, \dots, n\}$, the covariant derivative of a vector field $Y = Y^i \frac{\partial}{\partial x_i}$ in the direction of a vector $X = X^i \frac{\partial}{\partial x_i}$ is given by (Lee, 1997)

$$\nabla_X Y = \Big(\frac{\partial Y^i}{\partial x^j} X^j + \Gamma^i_{jk} X^j Y^k\Big) \frac{\partial}{\partial x^i}. \qquad (48)$$

The parallel displacement of the vector V_0 along the curve γ is the unique vector field $V : \mathbb{R} \to TM$ along γ satisfying $V(t_0) = V_0$ and whose covariant derivative is identically zero, $D_t V \equiv 0$. (This latter condition is sometimes also written, with a slight abuse of notation, $\nabla_{\dot\gamma} V \equiv 0$.) In coordinates, with $V(t) = V^i(t) \frac{\partial}{\partial x^i}$, the covariant derivative $D_t V$ is given by

$$(D_t V)(t) = \Big(\dot V^i(t) + \Gamma^i_{jk}(\gamma(t)) \dot\gamma(t)^j V^k(t)\Big) \frac{\partial}{\partial x^i},$$

where Γ^i_{jk} are the Christoffel symbols associated with the connection ∇. The condition $D_t V \equiv 0$ for parallel displacement thus gives the set of linear differential equations

$$\begin{aligned}\frac{d}{dt} V^i(t) &= -\Gamma^i_{jk}(\gamma(t)) \dot\gamma^j(t) V^k(t)\\ &=: \ A^i_k(\gamma(t), \dot\gamma(t))\, V^k(t), \quad i \in \{1, 2, \dots, n\} \end{aligned} \qquad (49)$$

Linearity with respect to $\dot\gamma(t)$ confirms that the parallel displacement is independent of the particular parametrization of the curve γ. In matrix form, (49) becomes

$$\dot v(t) = A(t) v(t), \qquad (50)$$

where the column vector $v(t) = (V^1; V^2; \dots; V^2)(t)$ is the vector representation of $V(t) \in T_{\gamma(t)}M$ in the chosen coordinate basis, with ; denoting row concatenation, and $A(t)$ the matrix with entries $A^i_j(\gamma(t), \dot\gamma(t))$, i, $j \in \{1, 2, \dots, n\}$. The solution of (50) with initial condition $v(t_0) = v_0$ (the coordinate representation of V_0) admits the expansion

$$v(t) = v_0 + (t - t_0) A(t_0) v_0 + o(t - t_0), \qquad (51)$$

providing a first order approximation to $P_\gamma^{t \leftarrow t_0} V_0$ in local coordinates.

Let ${}^1A$ and ${}^2A$ denote the time-varying matrices associated with ${}^1\nabla$ and ${}^2\nabla$, respectively, for a given set of local coordinates and a given curve γ as done in (49) above. Using (48) in the computation of covariant derivatives with ∇ given by the affine combination (46), we find that ∇ is indeed a connection and its Christoffel symbols are given by

$$\Gamma^i_{jk} = \alpha_1\, {}^1\Gamma^i_{jk} + \alpha_2\, {}^2\Gamma^i_{jk},$$

so that, by (49),

$$A(t) = \alpha_1\, {}^1A(t) + \alpha_2\, {}^2A(t). \qquad (52)$$

Also, note that ∇ can satisfy (48) *only* if $\alpha_1 + \alpha_2 = 1$ since otherwise the first term on the right hand side is not obtained. Using (52) in the local expression (51) for the parallel displacement associated with ∇ and noting that the parallel displacements for ${}^1\nabla$ and ${}^2\nabla$ also satisfy (51), i.e.,

$$\begin{aligned} {}^1v(t) &= v_0 + (t - t_0)\, {}^1A(t_0) v_0 + o(t - t_0),\\ {}^2v(t) &= v_0 + (t - t_0)\, {}^2A(t_0) v_0 + o(t - t_0), \end{aligned}$$

we conclude that (47) holds. □

An important application of this result is in the computation of the second covariant derivative of a map between manifolds using (1). In that expression, we take the limit as $t_1 \to t_0$ with t_0 fixed. Note however, that in one of the two parallel displacements employed, the varying t_1 is playing the role of the *initial* time. In this case, and as seen in (1), the argument of the parallel transport is in fact a vector field along γ rather than a fixed vector as indicated in the proposition.

Although the above result has been given for an affine combination of only two connections, it is easily generalized to an arbitrary finite number of connections.

The following corollary specializes the result to the parallel displacement with respect to the (0)-connection on Lie groups.

Corollary 5.2. Given a Lie group G and a smooth curve γ in G, the parallel displacement ${}^0P$ of a tangent vector v_0 along γ using the Cartan-Schouten (0)-connection satisfies

$$ {}^0P_\gamma^{t_1\leftarrow t_0} v_0 = \frac{1}{2}\big(\gamma_1\gamma_0^{-1}v_0 + v_0\,\gamma_0^{-1}\gamma_1\big) + o(t_1 - t_0) $$

with $\gamma_1 = \gamma(t_1)$, $\gamma_0 = \gamma(t_0)$, and $v_0 \in T_{\gamma_0}G$.

Proof: Denote the $(+)$, $(-)$, and (0) affine connections as ${}^+\nabla$, ${}^-\nabla$, and ${}^0\nabla$, respectively. Likewise, we use ${}^+P$, ${}^-P$, and ${}^0P$ for the associated parallel displacements.

The $(+)$ and $(-)$ Cartan-Schouten connections are flat and their path-independent parallel displacements satisfy (Mahony and Manton, 2002), respectively,

$$ {}^+P_\gamma^{t_1\leftarrow t_0} v_0 = v_0\gamma_0^{-1}\gamma_1 $$
$$ {}^-P_\gamma^{t_1\leftarrow t_0} v_0 = \gamma_1\gamma_0^{-1}v_0 $$

with γ_0, γ_1, and v_0 defined as in the statement of the corollary. Since the (0) affine connection satisfies

$$ {}^0\nabla = (1/2)\,{}^-\nabla + (1/2)\,{}^+\nabla $$

the result follows immediately from Proposition 5.1. $\square$

6. CONCLUSION

In this paper, we have provided a detailed discussion of the second covariant derivative of a map and its use in geometric optimal control calculations. Special care has been taken to fill in some of the historical holes that have been missing in our work to date. A detailed cheat sheet of the calculations needed to implement the projection operator on Lie groups have been compiled, exploiting in particular an efficient new operator notation. We have presented a new result providing a useful approximation for the parallel displacement of a connection that is an affine combination of connections for which the parallel displacements are known. This approximation is of special use in the calculation of the second covariant derivative of a map between Lie groups endowed with the symmetric Cartan-Schouten (0) connection. Together, the results described here provide a significant enabling technology for the numerical solution of optimal control problems on Lie groups. We believe that they can provide an interesting bridge between differential geometry and geometric optimal control.

7. ACKNOWLEDGEMENT

We would like to thank Robert Mahony and Jochen Trumpf for introducing us to the efficient operator approach and notation that has helped us in understanding and exposing these results.

We would also like to express our sincere gratitude to António M. Pascoal for his constant support, enthusiasm, and encouragement for our research. It is an honor to be amongst his collaborators and a privilege to enjoy his friendship.

REFERENCES

Anderson, B.D.O. and Moore, J.B. (1989). *Optimal Control: Linear Quadratic Methods*. Prentice-Hall.

Bryson, A.E. and Ho, Y.C. (1969). *Applied Optimal Control - Optimization, Estimation, and Control*. Blaisdell Publishing Company.

Eells, J. and Lemaire, L. (1978). A report on harmonic maps. *Bulletin of the London mathematical society*, 10(1978), 1–68.

Eells, J. and Sampson, J.H. (1964). Harmonic mappings of Riemannian manifolds. *American Journal of Mathematics*, 86(1), 109–160.

Ericksen, J.K. (1960). Tensor Fields. In Flügge (ed.), *Appendix to the Classical Field theories by C. Truesdell and R.A.Toupin in vol. III/1 of the Encyclopedia of Physics*, 794–850. Springer-Verlag, Berlin.

Hauser, J. and Meyer, D.G. (1998). The Trajectory Manifold of a Nonlinear Control System. In *37th IEEE Conference of Decision and Control (CDC)*, volume 1, 1034–1039.

Hauser, J. (2002). A Projection Operator Approach to the Optimization of Trajectory Functionals. In *Proceedings of the 15th IFAC World Congress*. Barcelona, Spain.

Hauser, J. (2003). On the Computation of Optimal State Transfers with application to the Control of Quantum Spin Systems'. In *Proceedings of the American Control Conference (ACC)*, 2169–2174. Denver, CO.

Hauser, J. and Saccon, A. (2006). A Barrier Function Method for the Optimization of Trajectory Functionals with Constraints. In *Proceedings of the 45th IEEE Conference on Decision and Control*, 864–869. IEEE.

Lee, J.M. (1997). *Riemannian manifolds: an introduction to curvature*. Springer, New York.

Mahony, R. and Manton, J.H. (2002). The Geometry of the Newton Method on Non-Compact Lie Groups. *Journal of Global Optimization*, 309–327.

Marsden, J.E. and Hughes, T.J.R. (1983). *Mathematical Foundations of Elasticity*. Prentice-Hall.

Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization*. 2nd edition.

Saccon, A., Aguiar, A.P., Häusler, A.J., Hauser, J., and Pascoal, A.M. (2012). Constrained Motion Planning for Multiple Vehicles on SE(3). In *IEEE Conference on Decision and Control (CDC)*, 3.

Saccon, A., Aguiar, A.P., and Hauser, J. (2011a). Lie Group Projection Operator Approach: Optimal Control on T SO(3). In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 6973–6978. Orlando, FL, USA.

Saccon, A., Hauser, J., and Aguiar, A.P. (2010). Optimal Control on Non-Compact Lie Groups: A Projection Operator Approach. In *49th IEEE Conference on Decision and Control (CDC)*, 7111–7116. IEEE.

Saccon, A., Hauser, J., and Aguiar, A.P. (2013). Optimal Control on Lie Groups: The Projection Operator Approach. *IEEE Transactions on Automatic Control (TAC), in press (available online)*.

Saccon, A., Hauser, J., and Aguiar, A. (2011b). Optimal Control on Lie Groups: Implementation Details of the Projection Operator Approach. In *IFAC World Congress*, 14567–14572.

Vilms, J. (1970). Totally geodesic maps. *Journal of Differential Geometry*, 4, 73–79.