Optimal Control on Non-Compact Lie Groups: A Projection Operator Approach

Alessandro Saccon, John Hauser and A. Pedro Aguiar

***Abstract*— Many nonlinear systems of practical interest evolve on Lie groups or on manifolds acted upon by Lie groups. Examples range from aircraft and underwater vehicles to quantum mechanical systems. In this paper, we develop the mathematical machinery needed for projection operator based trajectory exploration and optimization (optimal control) for systems defined on non-compact Lie groups.**

I. Introduction

The projection operator approach to the optimization of trajectory functionals, developed in [1], allows one to perform local Newton optimization of a (integral plus terminal) cost functional $h(\xi)$ over the Banach manifold $\mathcal{T}$ of trajectories of a nonlinear system (subject to a fixed initial condition). To work on the trajectory manifold, one *projects* curves ξ in the ambient Banach space onto the trajectory manifold, giving $\eta = \mathcal{P}(\xi) \in \mathcal{T}$, by using a local linear time-varying trajectory tracking controller. Noting that the constrained and unconstrained optimization problems

$$\min_{\xi \in \mathcal{T}} h(\xi) \quad \text{and} \quad \min_{\xi} h(\mathcal{P}(\xi))$$

are (essentially) locally equivalent [1, Section 2], one may develop Newton and quasi-Newton descent methods for trajectory optimization in an effectively unconstrained manner by working with the cost functional $\tilde{h}(\xi) = h(\mathcal{P}(\xi))$. In particular, the local Newton update, valid in a neighborhood of a second order sufficient local minimum, is given by $\xi_{i+1} = \mathcal{P}(\xi_i + \zeta_i)$ where

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

is the solution of a (finite horizon, time-varying) linear quadratic (LQ) optimal control problem. In the *flat* Banach space case, the usual chain rule applies and one finds that

$$\mathbf{D}^2\tilde{h}(\xi)\cdot(\zeta,\zeta) = \mathbf{D}^2h(\xi)\cdot(\zeta,\zeta) + \mathbf{D}h(\xi)\cdot\mathbf{D}^2\mathcal{P}(\xi)\cdot(\zeta,\zeta)$$

(for $\xi \in \mathcal{T}$, $\zeta \in T_\xi\mathcal{T}$) is a well defined object; see [2] for some projection operator calculus. The solution of the above LQ problem involves first and second order approximations of the nonlinear system about a given trajectory as well as the solution of some associated Riccati equations.

This work was supported in part by projects DENO/FCT-PT (PTDC/EEA-ACR/67020/2006), NAV-Control/FCT-PT (PTDC/EEA-ACR/65996/2006), Co3-AUVs (EU FP7 no. 231378), FCT-ISR/IST plurianual funding program, and the CMU-Portugal program. 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.

A. Saccon and A.P. Aguiar are with the Institute for Systems and Robotics (ISR), Instituto Superior Técnico (IST), Lisboa, Portugal {asaccon, pedro}@isr.ist.utl.pt

J. Hauser is with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, john.hauser@colorado.edu

When the system evolves on a Lie group, a number of interesting questions arise. What is the linearization of the system? How do we define and compute a second order approximation of the system? What Riccati equation(s) can we associate with a Lie group trajectory optimization problem? The purpose of this paper is to develop appropriate notions to address these questions.

II. Mathematical Preliminaries

In this section, we introduce standard definitions and notation that will be used thoughout the paper. We assume that the reader is familiar with the theory of finite dimensional smooth manifolds, matrix Lie groups, covariant differentiation. We refer to the books [3], [4], [5] for a review on differentiable manifolds and covariant differentiation and to [6], [7] for a review of the theory of Lie groups and Lie Algebra. A *smooth manifold* will be indicated with the letter M or N. A point on the manifold will be denoted simply by x. T_xM and T_x^*M denote, respectively, the *tangent* and *cotangent spaces* of M at x. A generic tangent vector is usually written as v_x or w_x, where the subscript indicates the base point at which the tangent vectors are attached. The *tangent* and *cotangent bundles* of M are denoted by TM and T^*M, respectively. The *natural bundle projection* from TM to M is denoted by $\pi : TM \to M$, so that $\pi\mathrm{v}_x = x$. A generic vector field on a manifold M is denoted by $X : M \to TM$. A vector field X is a *section* of the tangent bundle TM, that is it satisfies $\pi X(x) = x$. The *set of smooth vector fields* over M is denoted by $\mathfrak{X}(M)$.

Given a function $f : M \to N$, its *tangent map* is represented by $\mathbf{D}f : TM \to TN$ (or also as $Tf : TM \to TN$). Tangent maps act naturally on tangent vectors. Given a vector $\mathrm{v}_x \in T_xM$, $\mathbf{D}f(x) \cdot \mathrm{v}_x \in T_{f(x)}N$ (or $T_xf(\mathrm{v}_x)$) is the evaluation of the tangent map of f in the direction v_x at x. Tangent maps act naturally on vector fields as well. Given a vector field $X : M \to TM$, the writing $\mathbf{D}f \cdot X : M \to TN$ (and $Tf(X) : M \to TN$) denotes at $x \in M$ the tangent vector $\mathbf{D}f(x) \cdot X(x) \in T_{f(x)}N$ (and $T_xf(X(x))$). Given a diffeomorphism $\varphi : M \to N$ the *push-forward* of a vector field X on M through φ, denoted by φ_*X, is the vector field on N defined by $(\varphi_*X)(y) = T\varphi(X(\varphi^{-1}(y)))$, $y \in N$. Given a diffeomorphism $\varphi : M \to N$ the *pull-back* of a vector field Y on N through φ, written as φ^*Y, is the vector field on M defined by $(\varphi^*Y)(x) = T\varphi^{-1}(Y(\varphi(x)))$, $x \in M$, that is $(\varphi^{-1})_*Y$. Given an affine connection ∇ on a manifold M, we write ∇_XY and D_t to indicate respectively, the covariant derivative of the vector field Y in the direction X and the covariant differentiation with respect to the parameter t. The *parallel diplacement* along a curve $\gamma(t)$, $t \in I$, from $t = t_0$ to $t = t_1$ of a vector $V_0 \in T_{\gamma(t_0)}M$ is represented by $P_\gamma^{t_1 \leftarrow t_0} V_0$. We also adopt (see Section VI) the notation $\mathbb{D}Y \cdot X$ to mean $\nabla_X Y$. Given a function $f : M \to N$, $\mathbb{D}^2 f(x) \cdot (v_x, w_x) \in T_{f(x)}N$ is the second geometric derivative of f at $x \in M$ in the directions v_x, $w_x \in TM$ (see Section VIII).

A generic *Lie group* is denoted by G. The group *identity* is denoted by e. *Left* and *right translations* of $x \in G$ (a group element) by the group element $g \in G$ are denoted by $L_g x$ and $R_g x$, respectively. When convenient, we will adopt the shorthand notation gx, xg, $g\mathrm{v}_x$, $\mathrm{v}_x g$ for, in the same order, $L_g x$, $R_g x$, $T_x L_g(\mathrm{v}_x)$ and $T_x R_g(\mathrm{v}_x)$. A *left-invariant vector field* on G is a vector field such that $X(L_g x) = T_x L_g(X(x))$. Given a tangent vector at the identity $\varrho \in T_e G$, the symbol X_ϱ means the left-invariant vector field defined by $X_\varrho(g) := T_e L_g(\varrho)$. The *Lie algebra* of G is $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is identified with the tangent space $T_e G$ endowed with the *Lie bracket* operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, defined by $[\varrho, \varsigma] := [X_\varrho, X_\varsigma](e)$, where the latter bracket is the Jacobi-Lie bracket of the left-invariant vector fields X_ϱ and X_ς evaluated at the group identity.

The mapping $I_g(x) = gxg^{-1}$ is called *inner automorphism*. The *adjoint representation* of the Lie group G on the algebra $\mathfrak{g}$ is written as $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ and is the tangent map obtained differentiating $I_g(x)$ with respect to x at $x = e$. We recall that $\mathfrak{g} = T_e G$. Furthermore, the *adjoint representation* of the Lie algebra $\mathfrak{g}$ onto itself is written as $\mathrm{ad}_\varrho : \mathfrak{g} \to \mathfrak{g}$ and it is obtained differentiating $\mathrm{Ad}_g(\varsigma)$ with respect to g at $g = e$, in the direction ϱ. We recall that $\mathrm{ad}_\varrho \varsigma = [\varrho, \varsigma]$. The *exponential map* is denoted by $\exp : \mathfrak{g} \to G$ and its inverse (in a neighborhood of the identity) by $\log : G \to \mathfrak{g}$.

The trivialized tangent of $\log$ *and* $\exp$

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Consider a (local) diffeomorphism $F : G \to \mathfrak{g}$ between a neighborhood of the identity of G, N_e, and a neighborhood of the origin of $\mathfrak{g}$, N_0. Given $\xi \in N_0 \subseteq \mathfrak{g}$ the *(right) trivialized tangent* of F at ξ is the linear mapping $\mathbf{d}F_\xi : \mathfrak{g} \to \mathfrak{g}$ defined by

$$\mathbf{d}F_\xi \eta := \mathbf{D}F(g) \cdot TR_g \eta \,, \tag{1}$$

for $g = F^{-1}(\eta)$. Similarly, given $H : \mathfrak{g} \to G$, with $H = F^{-1}$, the *(right) trivialized tangent* of H at ξ is the linear mapping $\mathbf{d}H_\xi : \mathfrak{g} \to \mathfrak{g}$ defined by

$$\mathbf{d}H_\xi \eta := TR_{(H(\xi))^{-1}}(\mathbf{D}H(\xi) \cdot \eta) \tag{2}$$

More details on the trivialized tangent can be found in [8] and [9, section 4].

In this paper, we make use of the trivialized tangent of the logarithm and the exponential map, i.e., the particular case $F(g) := \log(g)$ and $H(\xi) := \exp(\xi)$.

III. Dynamical systems on Lie groups

A (smooth) control system on a Lie group G is a smooth mapping $f : G \times \mathbb{R}^m \times \mathbb{R} \to TG$, $(g, u, t) \mapsto f(g, u, t)$, such that $\pi f(g, u, t) = g$ for each $(g, u, t) \in G \times \mathbb{R}^m \times \mathbb{R}$.

A state trajectory of f is an absolutely continuous curve in G that satisfies (a.e.), for an assigned input $u(t)$,

$$\dot g(t) = f(g(t), u(t), t) \,. \tag{3}$$

Defining the *left trivialization* of f as $\lambda : G \times \mathbb{R}^m \times \mathbb{R} \to \mathfrak{g}$, $\lambda(g, u, t) := g^{-1} f(g, u, t)$, equation (3) can be equivalently written as

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

Remark. We would like to emphasize that the tangent bundle of a Lie group G *is* itself a Lie group, with operation $(g_1, v_1) \cdot (g_2, v_2) = (g_1 g_2, v_1 g_2 + g_1 v_2) \in T_{g_1 g_2} G$. This means that the theory developed in this paper is directly applicable, e.g., to the optimal control of mechanical systems evolving on Lie groups. ∎

Left-trivialized linearization around a trajectory

Let $\eta(t) = (g(t), u(t))$, $t \in [0, \infty)$, be a state-input trajectory of the control system (3). Given a bounded curve $v(t) \in \mathbb{R}^m$, $t \in [0, \infty)$, and $\varepsilon \in \mathbb{R}$ "small", consider the linear perturbation of the input defined as $u_\varepsilon(t) := u(t) + \varepsilon v(t)$. Indicating with g_ε the state trajectory associated with u_ε, we have

$$\dot g_\varepsilon(t) = g_\varepsilon(t) \lambda(g_\varepsilon(t), u_\varepsilon(t), t) \,, \tag{5}$$

$$g_\varepsilon(0) = g_0 \,. \tag{6}$$

In the (possibly small) interval $[0, T(\varepsilon))$, the solution g_ε will remain in a neighborhood of the unperturbed trajectory $g(t)$, $t \in [0, \infty)$, so that we can use the exponential coordinates to parameterize neighboring trajectories of the nominal state trajectory $g(t)$. To this end, we define the *left-trivialized perturbed trajectory* $z_\varepsilon(t)$, $t \in [0, T(\varepsilon))$, such that $g_\varepsilon(t) = g(t) \exp(z_\varepsilon(t))$, $t \in [0, T(\varepsilon))$. The trajectory z_ε satisfies the following differential equation.

Proposition 3.1: Let $x_\varepsilon(t) := \exp(z_\varepsilon(t))$, $t \in (0, T(\varepsilon))$. The *left trivialized perturbed trajectory* $z_\varepsilon(t)$, $t \in (0, T(\varepsilon))$, satisfies

$$\dot z_\varepsilon = \mathbf{d}\log_{z_\varepsilon} \Big(\mathrm{Ad}_{x_\varepsilon} \lambda\big(g x_\varepsilon, u_\varepsilon, t\big) - \lambda\big(g, u, t\big) \Big) \tag{7}$$

$$z_\varepsilon(0) = 0 \,. \tag{8}$$

Proposition 3.2 (Left-trivialized linearization): The left-trivialized perturbed trajectory $z_\varepsilon(t)$, $t \geq 0$, can be expanded to first order as $z_\varepsilon(t) = \varepsilon z(t) + R_2(\varepsilon, t)$ with R_2 of order higher that one in ε and $z(t)$ satisfying

$$\dot z(t) = A(\eta(t), t)\, z(t) + B(\eta(t), t)\, v(t) \,, \tag{9}$$

$$z(0) = z_0 \,, \tag{10}$$

where

$$A(\eta, t) := \mathbf{D}_1 \lambda(g, u, t) \circ TL_g - \mathrm{ad}_{\lambda(g,u,t)} \,, \tag{11}$$

$$B(\eta, t) := \mathbf{D}_2 \lambda(g, u, t) \,. \tag{12}$$

Definition 3.1: The *left trivialized linearization of* (3) *about the state-input trajectory* $\eta(t) \in G \times \mathbb{R}^m$*,* $t \geq 0$, is the linear system

$$\dot z(t) = A(\eta(t), t)\, z(t) + B(\eta(t), t)\, v(t) \,, \tag{13}$$

where $A(\eta(t), t)$ and $B(\eta(t), t)$ are given by (11) and (12), respectively.

IV. The Projection Operator

In this section we define the Projection Operator for systems evolving on a Lie Group. The Projection Operator is defined using a trajectory tracking controller that provides a convenient stable way to parameterize the neighborhood of a given state-input trajectory. We refer to [2, Section 1] for an introduction of the notion of Projection Operator for nonlinear systems evolving on $\mathbb{R}^n$.

Let $f : G \times \mathbb{R}^m \to TG$ be a (time-invariant) control system. A state-input trajectory $\xi = (\alpha, \mu)$ of f is called *exponentially stabilizable* if (and only if) there is a feedback law $u(t) = k(g(t), \alpha(t), \mu(t), t)$, with $k(\alpha(t), \alpha(t), \mu(t), t) = \mu(t)$ for all $t \geq 0$, such that α is an exponentially stable (state) trajectory of the closed loop system

$$\dot{g}(t) = f\big(g(t), k\big(g(t), \alpha(t), \mu(t), t\big)\big) , \quad g(0) = g_0 , \quad (14)$$

that is, there exist $M < \infty$, $\lambda > 0$, and $\delta > 0$ such that $\| \log(g(t)^{-1}\alpha(t))\| \leq M e^{-\lambda t} \| \log(g_0^{-1}\alpha(0))\|$ for all $t \geq 0$ and all g_0 in a neighborhood of $\alpha(0)$ such that $\| \log(g_0^{-1}\alpha(0))\| < \delta$.

We would also impose some smoothness and boundedness conditions on k. We may restrict our attention to feedback of the form

$$\begin{aligned} u(t) &= k(g(t), \alpha(t), \mu(t), t) \\ &= \mu(t) + K(t)\big[\log(g(t)^{-1}\alpha(t))\big] , \end{aligned} \quad (15)$$

as a trajectory ξ of a C^1 nonlinear system is exponentially stabilizable if and only if there is a bounded gain matrix K that stabilizes the (left-trivialized) linearization of f about ξ. It will be evident from next section that the linearization of (14) with feedback (15) around a state trajectory α is given by the linear system (without input)

$$\dot{z}(t) = [A(\xi(t), t) - B(\xi(t), t)K(t)]z(t) . \quad (16)$$

Definition 4.1 (Projection Operator $\mathcal{P}$): For a given initial condition $g_0 \in G$ and a time-varying feedback $K(t)$, $t \geq 0$, equation (14) with feedback (15) defines a causal operator, called the *Projection Operator*, which maps a given curve $\xi(t) = (\alpha(t), \mu(t)) \in G \times \mathbb{R}^m$, $t \geq 0$, into the state-input trajectory $\eta(t) = (g(t), u(t)) \in G \times \mathbb{R}^m$, $t \geq 0$, that satisfies

$$\dot{g} = g\lambda_K(g, \xi(t), t) , \quad (17)$$

$$u(t) = u_K(g, \xi(t), t)) , \quad (18)$$

$$g(0) = g_0 , \quad (19)$$

where

$$\lambda_K(g, \xi, t) := \lambda(g, u_K(g, \xi, t)) , \quad (20)$$

$$u_K(g, \xi, t) := \mu + K(t) \log(g^{-1}\alpha) . \quad (21)$$

In short, we write $\eta = \mathcal{P}_K^{g_0}(\xi)$ or, when g_0 and K are clear from the context, simply $\eta = \mathcal{P}(\xi)$. The Projection Operator satisfies the projection property $\mathcal{P}(\xi) = \mathcal{P}(\mathcal{P}(\xi)) =: \mathcal{P}^2(\xi)$.

V. The first order approximation of the Projection Operator

Let $\xi(t) = (\alpha(t), \mu(t))$, $t \geq 0$, be a curve in $G \times \mathbb{R}^m$ and $\zeta(t) = (\beta(t), \nu(t))$, $t \geq 0$, a curve in $\mathfrak{g} \times \mathbb{R}^m$. In the following, we write $\exp(\zeta)$ and $\log(\xi)$ for the point-wise operators defined by posing $\exp(\zeta)(t) = (\exp(\beta(t)), \nu(t)) \in G \times \mathbb{R}^m$ and $\log(\xi)(t) = (\log(\alpha(t)), \mu(t)) \in G \times \mathbb{R}^m$, $t \geq 0$.

We are interested in studying the effect of a perturbation of the curve ξ in the direction ζ, that is to study the mapping $\mathcal{P}(\xi \exp(\varepsilon\zeta))$, for $\varepsilon \in \mathbb{R}$ "small". From continuity of the mapping $\mathcal{P}$, we can parameterize $\mathcal{P}(\xi \exp(\varepsilon\zeta))$ using the left-trivialized perturbed trajectory $\chi_\varepsilon(t) \in \mathfrak{g} \times \mathbb{R}^m$, $t \geq 0$, defined by

$$\mathcal{P}(\xi \exp(\varepsilon\zeta)) = \mathcal{P}(\xi) \exp(\chi_\varepsilon) . \quad (22)$$

Definition 5.1: The left-trivialized *Local Projection Operator* around the curve ξ, that we write $\chi = \mathcal{N}_\xi(\zeta)$, is the operator that takes a curve $\zeta(t) = (\beta(t), \nu(t)) \in \mathfrak{g} \times \mathbb{R}^m$, $t \geq 0$, to the left-trivialized trajectory $\chi(t) = (y(t), w(t)) \in \mathfrak{g} \times \mathbb{R}^m$, $t \geq 0$, defined by

$$\mathcal{N}_\xi(\zeta) := \log(\mathcal{P}(\xi)^{-1}\mathcal{P}(\xi \exp(\zeta))) . \quad (23)$$

Proposition 5.1: Given the curves $\xi = (\alpha, \mu)$ and $\eta = (g, u)$ such that $\eta = P(\xi)$, the mapping $(y_\varepsilon, w_\varepsilon) = \chi_\varepsilon = \mathcal{N}_\xi(\varepsilon\zeta) = \mathcal{N}_\xi(\varepsilon\beta, \varepsilon\nu)$ can be computed explicitly as

$$\begin{aligned} \dot{y}_\varepsilon = \mathbf{d}\log_{y_\varepsilon} \big[&\mathrm{Ad}_{\exp y_\varepsilon} \\ &\lambda_K(g \exp y_\varepsilon, \xi \exp(\varepsilon\zeta), t) - \lambda_K(g, \xi, t)\big] , \end{aligned} \quad (24)$$

$$w_\varepsilon(t) = u_K(g \exp y_\varepsilon, \xi \exp(\varepsilon\zeta), t) - u_K(g, \xi, t) , \quad (25)$$

$$y_\varepsilon(0) = 0 . \quad (26)$$

Proposition 5.2: The left-trivialized trajectory $\chi_\varepsilon = \mathcal{N}_\xi(\varepsilon\zeta)$ can be expanded to first order as $\chi_\varepsilon(t) = \varepsilon\gamma(t) + R_2(\varepsilon, t)$ with R_2 of order higher than one in ε and $\gamma(t) = (z(t), v(t))$, $t \geq 0$, satisfies

$$\gamma = \mathbf{D}\mathcal{N}_\xi(0) \cdot \zeta = \mathcal{P}(\xi)^{-1}\mathbf{D}\mathcal{P}(\xi) \cdot \xi\zeta \quad (27)$$

and can be computed using

$$\dot{z} = A(\eta(t))\, z + B(\eta(t))\, v , \quad (28)$$

$$v = \nu + K\mathbf{d}\log_{\log(g^{-1}\alpha)}(\mathrm{Ad}_{g^{-1}\alpha}\beta - z) , \quad (29)$$

$$z(0) = 0 , \quad (30)$$

where $A(\eta(t))$ and $B(\eta(t))$ are given by (11) and (12). Note that we have dropped the time dependence of the matrices A and B as the (trivialized) vector field λ is time invariant.

The proof is based on perturbation theory much as Proposition (3.2) is.

VI. Affine connections, covariant derivative, and parallel displacement along a curve

We recall in this section the basic properties and facts about affine connections, covariant derivative, and parallel displacement along a curve, referring to, e.g., [3, Chapter 7] and [5, Chapter 4] for further details. This section should be seen as a preliminary material that will allow us to define the concept of second *geometric derivative* that we introduce in Section VIII to compute, in Section XI, the second order approximation of the Projection Operator $\mathcal{P}$.

Definition 6.1: An *affine connection* or *covariant derivative* ∇ (pronounced "del" or "nabla") on a smooth manifold M is a map $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ written as $(X, Y) \mapsto \nabla_X Y$ that satisfies the following properties

1) $\nabla_X Y$ is linear over $C^\infty(M)$ in X

$$\nabla_{fX_1+gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y\,,\ f,g \in C^\infty(M)\,;$$

2) $\nabla_X Y$ is linear over $\mathbb{R}$ in Y

$$\nabla_X(aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2\,,\ a,b \in \mathbb{R}\,;$$

3) $\nabla_X Y$ satisfies the following derivation rule

$$\nabla_X(fY) = (Xf)\,Y + f\,\nabla_X Y\,,\quad f \in C^\infty(M)\,.$$

Having a connection at hand, one can now define the concept of *covariant derivative along a curve*. Let I be a opened interval on $\mathbb{R}$. Given a curve $\gamma : I \to M$, we denote by $\mathcal{T}(\gamma)$ the space of smooth functions $I \to TM|_\gamma$, such that if $V \in \mathcal{T}(\gamma)$ then $V(t) \in T_{\gamma(t)}M$, $t \in I$.

Proposition 6.1: Let ∇ be a linear connection on M. For each curve $\gamma : I \to M$, ∇ determines a unique operator

$$D_t : \mathcal{T}(\gamma) \to \mathcal{T}(\gamma)$$

satisfying for each V and $W \in \mathcal{T}(\gamma)$

1) Linearity over $\mathbb{R}$

$$D_t(aV + bW) = a\,D_t V + b\,D_t W\,,\ a,b \in \mathbb{R}\,;$$

2) Product rule

$$D_t(fV) = \dot{f}\,V + f\,D_t V\,,\quad f \in C^\infty\,;$$

3) If $V \in \mathcal{T}(\gamma)$ is *extendible* outside γ so that $V(t) = Y(\gamma(t))$, $Y \in \mathfrak{X}(M)$, then

$$D_t V(t) = \big(\nabla_{\dot\gamma(t)} Y\big)(\gamma(t))\,.$$

The covariant derivative allow us to define the concept of *parallel displacement of a vector along a curve*.

Definition 6.2: Let M be a smooth manifold endowed with an affine connection ∇. Given a curve $\gamma : I \to M$, the *parallel displacement* of the vector $V_0 \in T_{\gamma(t_0)}M$, $t_0 \in I$, along γ at time $t_1 \in I$ is given by the *unique* vector field V along γ, $V \in \mathcal{T}(\gamma)$, that satisfies

$$\begin{aligned} D_t V(t) &= 0\,,\quad t \in I\,,\\ V(t_0) &= V_0\,, \end{aligned} \tag{31}$$

evaluated at time t_1. We will denote such a (linear) transformation by $P_\gamma^{t_1\leftarrow t_0} V_0 := V(t_1)$.

Note that given vector fields X_1, X_2, and Y with $X_1(x) = X_2(x)$ one has $(\nabla_{X_1} Y)(x) = (\nabla_{X_2} Y)(x)$. For this reason, and others that will be discussed in Section VIII, we will equivalently write covariant differentiation as $\mathbb{D}Y(x)\cdot X(x)$ in place of $(\nabla_X Y)(x)$. In this fashion, e.g., $(D_t)\,Y(\gamma(t)) = \mathbb{D}Y(\gamma(t))\cdot\dot\gamma(t)$.

VII. Bi-invariant affine connections on Lie groups

Affine connections, covariant differentiation, and parallel displacement can be defined on an arbitrary Lie group since each Lie group has a smooth manifold structure. Furthermore, the tangent bundle of a Lie group is trivial as we can always identify TG with the product $G \times \mathfrak{g}$. Amongst all possible affine connections ∇, *left-invariant* connections are those which commute with the push-forward of the left translation, that is

$$(L_g)_* \nabla_X Y = \nabla_{(L_g)_* X}(L_g)_* Y\,, \tag{32}$$

while *right-invariant* connections are those which commute with the push-forward of the right translation. A connection is *bi-invariant* if it is both right- and left-invariant. We now recall the following result (see [11, Theorem 8.1]).

Lemma 7.1: The G be a Lie group. There is a one-to-one correspondence between left-invariant (*resp., right-invariant*) affine connections on G and bilinear maps $\omega : \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ given by

$$\omega(\varrho,\varsigma) = \big(\nabla_{X_\varrho} X_\varsigma\big)(e)\,,\quad \varrho,\varsigma \in \mathfrak{g} \tag{33}$$

for $X_\varrho(g) := TL_g\varrho$ (*resp.,* $:= TR_g\varrho$).

The bilinear function ω appearing in (33) is termed the *left* (respectively, *right*) *connection function* for ∇. A connection function is strictly related to the Christoffel's symbols defining the connection. However, an invariant connection is uniquely specified by assigning n^3 *numbers*, i.e., the bilinear connection function, as opposed to n^3 *functions*, i.e., the Christoffel's symbols (where n is the manifold dimension).

The $(+)$, $(-)$, and (0) Cartan-Schouten affine connections on connected Lie Groups

Amongst all possible bi-invariant affine connections, three are particularly useful: they are the (0), (+) and (-) Cartan-Schouten connections. These connections were studied and generalized to homogeneous spaces by Nomizu in [11, Section 11], although in the context of Lie groups they were introduced by E. Cartan in [12].

One extremely useful property of these bi-invariant connections is that the parallel displacement of a vector along a curve is *independent* of the path, depending only on the initial and final points of the curve. Also, every 1-parameter subgroup $\gamma_\varrho(t) := \exp(t\varrho)$ is a *geodesic*, that is, $\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0$. For bi-invariant connections, the right and left-connection functions coincide and we have $\omega(\mathrm{Ad}_x\varrho, \mathrm{Ad}_x\varsigma) = \mathrm{Ad}_x\omega(\varrho,\varsigma)$. See [11, Section 11].

Let $\gamma : I \to G$ be a curve on G such that $\gamma(t_0) = x_0$ and $\gamma(t_1) = x_1$, for t_0, $t_1 \in I$. We have

- The $(-)$ connection satisfies

$$\omega(\varrho,\varsigma) = 0\,,\qquad \varrho,\varsigma \in \mathfrak{g}\,, \tag{34}$$

$$P_\gamma^{t_1\leftarrow t_0}\mathrm{v}_0 = x_1 x_0^{-1}\mathrm{v}_0\,,\qquad v_0 \in T_{x_0}G\,; \tag{35}$$

- The $(+)$ connection satisfies

$$\omega(\varrho,\varsigma) = [\varrho,\varsigma]\,,\qquad \varrho,\varsigma \in \mathfrak{g}\,, \tag{36}$$

$$P_\gamma^{t_1\leftarrow t_0}\mathrm{v}_0 = \mathrm{v}_0\, x_0^{-1} x_1\,,\qquad v_0 \in T_{x_0}G\,; \tag{37}$$

- The (0) connection satisfies

$$\omega(\varrho,\varsigma) = 1/2\,[\varrho,\varsigma]\,,\qquad \varrho,\varsigma \in \mathfrak{g}\,, \tag{38}$$

$$P_\gamma^{t_1\leftarrow t_0}\mathrm{v}_0 = 1/2\,(\,x_1 x_0^{-1}\mathrm{v}_0 + \mathrm{v}_0 x_0^{-1} x_1)\,,\ v_0 \in T_{x_0}G\,. \tag{39}$$

Note that the (0) connection is obtained as the arithmetic mean of the $(-)$ and $(+)$ connections. (An affine combination of affine connections is always an affine connection.)

VIII. The geometric derivative: covariant differentiation of maps between manifolds

Let M_1 and M_2 be two smooth manifolds endowed with affine connections ${}^1\nabla$ and ${}^2\nabla$, respectively, and let $f : M_1 \to M_2$ be a smooth mapping. The second geometric derivative is a tool to extend the classical (Leibniz's) product rule to the covariant derivative of the "product" $Df(\gamma_1(t)) \cdot V_1(t)$, for a curve γ_1 and a vector field V_1 along γ_1 in M_1.

Chosen $x \in M_1$ and two tangent vectors v_x and $\mathrm{w}_x \in T_xM_1$, let $\gamma_1 : I \to M_1$ be a smooth curve in M_1 such that $\gamma_1(t_0) = x$ and $\dot\gamma_1(t_0) = \mathrm{w}_x$, V_1 a smooth vector field along γ_1 such that $V_1(t_0) = \mathrm{v}_x$, and $V_2(t) := \mathbf{D}f(\gamma_1(t)) \cdot V_1(t) \in T_{f(\gamma_1(t))}M_2$ a smooth vector field along the curve $\gamma_2(t) := f(\gamma_1(t))$ in M_2.

Definition 8.1: The *second geometric derivative* of the map $f : M_1 \to M_2$ at $x \in M_1$ in the directions v_x and $\mathrm{w}_x \in T_xM_1$ is the bilinear mapping $\mathbb{D}^2 f(x) : T_xM_1 \times T_xM_1 \to T_{f(x)}M_2$ defined as

$$\mathbb{D}^2 f(x) \cdot (\mathrm{v}_x, \mathrm{w}_x) = D_t V_2(t_0) - \mathbf{D}f(\gamma_1(t_0)) \cdot D_t V_1(t_0)\,, \quad (40)$$

where D_tV_1 and D_tV_2 denote the covariant differentiation with respect to ${}^1\nabla$ and ${}^2\nabla$, respectively.

Corollary 8.1: Denote by ${}^1P$ and ${}^2P$ the parallel displacements associated to ${}^1\nabla$ and ${}^2\nabla$, respectively. Then, equation (40) is equal (for $t = t_0$) to

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\Big({}^2P_{\gamma_2}^{t\leftarrow t+\varepsilon}\, \mathbf{D}f(\gamma_1(t+\varepsilon)) \cdot {}^1P_{\gamma_1}^{t+\varepsilon\leftarrow t} V_1(t) - \mathbf{D}f(\gamma_1(t)) \cdot V_1(t)\Big)\,, \quad (41)$$

Proof: The connection ${}^2\nabla$ allows us to compute the covariant derivative of the vector field V_2 along γ_2 as

$$\big(D_tV_2\big)(t) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left({}^2P_{\gamma_2}^{t\leftarrow t+\varepsilon} V_2(t+\varepsilon) - V_2(t)\right)\,, \quad (42)$$

Equation (42) can be expanded into

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\Big({}^2P_{\gamma_2}^{t\leftarrow t+\varepsilon}\mathbf{D}f(\gamma_1(t+\varepsilon)) \cdot V_1(t+\varepsilon) - \mathbf{D}f(\gamma_1(t)) \cdot V_1(t)\Big)\,. \quad (43)$$

Adding and subtracting the term ${}^2P_{\gamma_2}^{t\leftarrow t+\varepsilon}\, \mathbf{D}f(\gamma_1(t+\varepsilon)) \cdot {}^1P_{\gamma_1}^{t+\varepsilon\leftarrow t} V_1(t)$ inside the parenthesis of the previous expression, and noting that (in TM_2)

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon} {}^2P_{\gamma_2}^{t\leftarrow t+\varepsilon}\mathbf{D}f(\gamma_1(t+\varepsilon)) \cdot \big(V_1(t+\varepsilon) - {}^1P_{\gamma_1}^{t+\varepsilon\leftarrow t} V_1(t)\big)$$

is equal to $Df(\gamma_1(t)) \cdot D_tV_1(t)$, the result follows. ■

Remark. The *second geometric derivative* defined above does not appear to be a standard notion in differential and Riemannian geometry. When $M_1 = M_2$ and $f = \mathrm{id}$ (the identity map), $\mathbb{D}^2 f$ reduces to the difference of the connections ${}^2\nabla$ and ${}^1\nabla$, that is $\mathbb{D}^2\mathrm{id}(x) \cdot (X(x), Y(x)) = ({}^2\nabla_X Y - {}^1\nabla_X Y)(x)$, which is known to be a $(1, 2)$-tensor. Also, the second geometric derivative reduces to the *second covariant derivative* when $M_2 = \mathbb{R}$ (that is, f is a real function). We refer to [13, Sections 5.6 and 5.7] and reference therein, for further reading. In the context of Riemannian geometry, we suspect that the second geometric derivative is strictly related to the concept of *second fundamental form* when $f : M_1 \to M_2$ is an isometric embedding of M_1 into M_2. ■

Note that one can define the concept of higher order geometric derivative ($\mathbb{D}^3 f$, $\mathbb{D}^4 f$, and so on) by imposing the product rule to hold for the covariant differentiation. For reason of space, we will not present explicit formulae here, although we will make use of $\mathbb{D}^3 \log$ in computing the second order approximation of $\mathcal{N}_\xi$ in Section XI.

IX. The geometric derivative on Lie groups: Differentiation rules for the (0) connection

The second (and higher order) geometric derivative(s) of a function between two Lie groups can be computed as soon as we specify affine connections on domain and codomain.

In this section, we restrict our attention to second geometric derivatives with respect to the (0)-connection, as this will be the one we will use for computing the second order approximation of the Projection Operator in Section XI. One can prove the following.

Proposition 9.1: Let $X(g) := g\xi(g)$ and $Y(g) = \xi(g)g$ be two vector fields on G with $g \mapsto \xi(g)$ a $\mathfrak{g}$-valued function. The following holds.

$$\begin{aligned} \mathbb{D}X(g) \cdot g\eta &= g\big(\mathbf{D}\xi(g) \cdot g\eta + 1/2\,[\eta, \xi(g)]\big)\,, \\ \mathbb{D}Y(g) \cdot \eta g &= \big(\mathbf{D}\xi(g) \cdot \eta g + 1/2\,[\xi(g), \eta]\big)g\,. \end{aligned} \quad (44)$$

Proposition 9.2: Let $G \ni g \mapsto \xi(g) \in \mathfrak{g}$ be defined as $g^{-1}X(g) = T_gL_{g^{-1}}X(g)$, with $X(g)$ a vector field. Then

$$\mathbf{D}\xi(g) \cdot g\eta = g^{-1}\Big(\mathbb{D}X(g) \cdot g\eta\Big) + 1/2\,\big[g^{-1}X(g), \eta\big]\,.$$

Proposition 9.3: For each $\varrho, \varsigma \in \mathfrak{g}$, we have

$$\mathbb{D}^2 \exp(0) \cdot (\varrho, \varsigma) = 0\,, \qquad \mathbb{D}^2 \log(e) \cdot (\varrho, \varsigma) = 0\,.$$

Proposition 9.4: Let $t \mapsto V(t)$ be a vector field along the curve $\gamma \in G$. Let $W(t) := g\,V(t)$ (W is a vector field along the curve $g\gamma$). Then $D_tW(t) = g\,D_tV(t)$.

X. Quadratic approximation of the unconstrained problem

Consider the cost functional

$$h(\xi) := \int_0^{t_f} l(\xi(\tau), \tau)\, d\tau + m(\pi_1\xi(t_f))\,, \quad (45)$$

where $\xi(t) = (\alpha(t), \mu(t))$ and $\pi_1\xi(t) = \alpha(t)$. Using the Projection Operator $\mathcal{P}$, the functional $\tilde h$ over the space of curves in $G \times \mathbb{R}^m$ is constructed as

$$\tilde h(\xi) = h(\mathcal{P}(\xi))\,. \quad (46)$$

As mentioned in the introduction, our goal is to find a quadratic approximation of $\tilde h$ around a given curve ξ. To this end, recalling the definition of the (left-trivialized) Local Projection operator $\mathcal{N}_\xi$ (Def. 5.1) we obtain, for a given perturbation $\zeta(t) \in \mathfrak{g} \times \mathbb{R}^m$, $t \in \mathbb{R}$, the identity

$$h\big(\mathcal{P}(\xi \exp \varepsilon\zeta)\big) = h\big(\mathcal{P}(\xi) \exp(\mathcal{N}_\xi(\varepsilon\zeta))\big)\,. \quad (47)$$

Note that the above expression, as a function of ε and for fixed ξ and ζ, defines a real function on $\mathbb{R}$. Expanding the left hand side of (47) with respect to ε, using the fact that $\mathbf{D}\exp(0)\cdot\zeta = \zeta$ and $\mathbb{D}^2\exp(0) = 0$, gives

$$h(\mathcal{P}(\xi\exp(\varepsilon\zeta))) = h(\mathcal{P}(\xi)) + \varepsilon\,\mathbf{D}h(\mathcal{P}(\xi))\cdot\mathbf{D}\mathcal{P}(\xi)\cdot\xi\zeta \\ + 1/2\,\varepsilon^2\left[\mathbb{D}^2 h(\mathcal{P}(\xi))\cdot(\mathbf{D}\mathcal{P}(\xi)\cdot\xi\zeta, \mathbf{D}\mathcal{P}(\xi)\cdot\xi\zeta)\right. \\ \left. + \mathbf{D}h(\mathcal{P}(\xi))\cdot\mathbb{D}^2\mathcal{P}(\xi)\cdot(\xi\zeta,\xi\zeta)\right] + o(\varepsilon^2) \quad (48)$$

where the first and second geometric derivative of h are

$$\mathbf{D}h(\xi)\cdot\xi\zeta = \int_0^{t_f}\mathbf{D}l(\xi(\tau),\tau)\cdot\xi(\tau)\zeta(\tau)\,d\tau \\ + \mathbf{D}m(\pi_1\xi(t_f))\cdot T\pi_1(\xi(t_f)\zeta(t_f)) \quad (49)$$

and

$$\mathbb{D}^2 h(\xi)\cdot(\xi\zeta_1,\xi\zeta_2) = \quad (50) \\ \int_0^{t_f}\mathbb{D}^2 l(\xi(\tau),\tau)\cdot(\xi(\tau)\zeta_1(\tau),\xi(\tau)\zeta_2(\tau))\,d\tau \\ + \mathbb{D}^2 m(\pi_1\xi(t_f))\cdot\big(T\pi_1(\xi(t_f)\zeta(t_f)), T\pi_1(\xi(t_f)\zeta(t_f))\big)$$

with $T\pi_1(\xi(t)\zeta(t)) = \alpha(t)\beta(t)$. Expanding the right hand side of (47) with respect to ε, we get

$$h(\mathcal{P}(\xi)\exp(\mathcal{N}_\xi(\varepsilon\zeta))) = h(\mathcal{P}(\xi)) \\ + \varepsilon\,\mathbf{D}h(\mathcal{P}(\xi))\cdot\mathcal{P}(\xi)\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta \\ + 1/2\,\varepsilon^2\big[\mathbb{D}^2 h(\mathcal{P}(\xi))\cdot(\mathcal{P}(\xi)\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta, \mathcal{P}(\xi)\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta) \\ + \mathbf{D}h(\mathcal{P}(\xi))\cdot\mathcal{P}(\xi)\big(\mathbb{D}^2\exp(0)\cdot(\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta,\ \mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta) \\ + \mathbf{D}^2\mathcal{N}_\xi(0)\cdot(\zeta,\zeta)\big)\big] + o(\varepsilon^2)\,. \quad (51)$$

From the definition of $\mathcal{N}_\xi$, we get

$$\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta_1 = \mathcal{P}(\xi)^{-1}\mathbf{D}\mathcal{P}(\xi)\cdot\xi\zeta_1 \quad (52)$$

and, furthermore,

$$\mathbf{D}^2\mathcal{N}_\xi(0)\cdot(\zeta_1,\zeta_2) = \mathcal{P}(\xi)^{-1}\mathbb{D}\mathcal{P}^2(\xi)\cdot(\xi\zeta_1,\xi\zeta_2) \\ + \mathbb{D}^2\log(e)\cdot(\mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta_1, \mathbf{D}\mathcal{N}_\xi(0)\cdot\zeta_2)\,. \quad (53)$$

As mentioned in Section IX, using the (0) connection $\mathbb{D}^2\exp(0)\cdot(\zeta_1,\zeta_2) = 0$ and $\mathbb{D}^2\log(e)\cdot(\zeta_1,\zeta_2) = 0$. Therefore, we obtain the following result.

Proposition 10.1: The second geometric derivative of the $\mathcal{P}$ with respect to the (0) connection satisfies

$$\mathbf{D}^2\mathcal{N}_\xi(0)\cdot(\zeta_1,\zeta_2) = \mathcal{P}(\xi)^{-1}\mathbb{D}^2\mathcal{P}(\xi)\cdot(\xi\zeta_1,\xi\zeta_2)\,. \quad (54)$$

Note that we write $\mathbf{D}^2\mathcal{N}_\xi$ instead of $\mathbb{D}^2\mathcal{N}_\xi$ to highlight the fact that $\mathcal{N}_\xi$ is an operator between two Euclidean spaces.

It is quite instructive to compare (48) with (51): As those expressions express the same quantity, one can guess (and then prove!) that $\mathbb{D}^2\exp(0)\cdot(\varrho,\varsigma)$ = - $\mathbb{D}^2\log(e)\cdot(\varrho,\varsigma)$, $\forall\varrho,\varsigma\in\mathfrak{g}$, *for any choice* of connection!

Summarizing the above discussion, we can conclude that the quadratic expansion of $\tilde{h}$ is as follows.

Proposition 10.2: The expansion with respect to ε of the functional $\tilde{h}(\xi\exp(\varepsilon\zeta))$ is given by (48), where the derivatives $\mathbf{D}\mathcal{P}(\xi)\cdot\xi\zeta$ and $\mathbb{D}^2\mathcal{P}(\xi)\cdot(\xi\zeta_1,\xi\zeta_2)$, with (0) connection, can be computed using (52) and (54), respectively.

XI. The second derivative of $\mathcal{N}_\xi$

This section shows the relationship between the second derivative of the Local Projection Operator $\mathcal{N}$ and the second geometric derivative of the Projection Operator $\mathcal{P}$ at ξ.

Proposition 11.1: Given a trajectory $\eta = (g,u)$, $\eta = \mathcal{P}(\eta)$, the second derivative of the (left-trivialized) Local Projection Operator $\mathcal{N}_\xi$ around zero in the directions ζ_1 and ζ_2,

$$(y,w) = \mathbf{D}^2\mathcal{N}_{(g,u)}(0)\cdot((\beta_1,\nu_1),(\beta_2,\nu_2)) \\ = \mathbf{D}^2\mathcal{N}_\eta(0)\cdot(\zeta_1,\zeta_2) = \mathcal{P}(\eta)^{-1}\mathbb{D}^2\mathcal{P}(\eta)\cdot(\eta\zeta_1,\eta\zeta_2)\,,$$

is given by

$$\dot{y} = A(\eta)y + B(\eta)w - 1/2\left((\mathrm{ad}_{z_1}\mathrm{ad}_{z_2} + \mathrm{ad}_{z_2}\mathrm{ad}_{z_1})\lambda(\eta)\right. \\ \left. - \mathrm{ad}_{z_1}(A(\eta)z_2 + B(\eta)v_2) - \mathrm{ad}_{z_2}(A(\eta)z_1 + B(\eta)v_1)\right) \\ + \mathbb{D}^2\lambda(\eta)\cdot(\eta\gamma_1,\eta\gamma_2)\,, \quad (55)$$

$$w = -K(t)\big[y + 1/2\left([z_1,\beta_2] + [z_2,\beta_1]\right)\big]\,, \quad (56)$$

with $y(0) = 0$, $\gamma_i = (z_i,v_i) = D\mathcal{N}_\eta(0)\cdot\zeta_i$, $i = \{1,2\}$, and where $A(\eta)$ and $B(\eta)$ are defined in (11) and (12), respectively. Note that for brevity we have suppressed the presence of the t argument in expressions (55) and (56).

XII. Conclusion

In this paper, we have extended the projection operator based trajectory optimization approach to the class of non-linear systems that evolve on non-compact Lie groups. This required the introduction of a geometric derivative notion for the repeated differentiation of a mapping between two Lie groups, endowed with affine connections. With this tool, chain rule like formulas where used to develop expressions for the basic objects needed for trajectory optimization.

We thank the anonymous reviewers for their valuable comments.

References

[1] J. Hauser, "A projection operator approach to the optimization of trajectory functionals," in *15th IFAC World Congress*, Barcelona, Spain, 2002.
[2] J. Hauser and D. Meyer, "The trajectory manifold of a nonlinear control system," in *37th IEEE Conference of Decision and Control (CDC)*, vol. 1, 1998, pp. 1034–1039.
[3] W. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, 2nd ed., ser. Pure and applied mathematics. Academic Press, Boston, 1986.
[4] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*. Springer-Verlag, New York, 1988.
[5] J. M. Lee, *Riemannian manifolds: an introduction to curvature*. Springer, New York, 1997.
[6] V. Varadarajan, *Lie groups, Lie algebras, and their representations*. Springer-Verlag, New York, 1984.
[7] W. Rossmann, *Lie groups. an introduction through linear groups*. Oxford University Press, 2002.
[8] A. Iserles, H. Munthe-Kaas, S. Nørsett, and A. Zanna, "Lie-group methods," *Acta Numerica*, vol. 9, pp. 1–148, 2000.
[9] N. Bou-Rabee and J. Marsden, "Hamilton-pontryagin integrators on lie groups part i: Introduction and structure-preserving properties," *Foundations of Computational Mathematics (FOCM)*, vol. 9, no. 2, pp. 197–219, April 2009.
[10] H. Khalil, *Nonlinear Systems*, 2nd ed. Prentice Hall, 1996.
[11] K. Nomizu, "Invariant affine connections on homogeneous spaces," *American Journal of Mathematics*, vol. 76, pp. 33–65, 1954.
[12] E. Cartan, "La géométrie des groupes de transformations," *Journal de Mathématiques pures et appliquées*, vol. 6, no. 9, pp. 1–119, 1927.
[13] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on Matrix Manifolds*. Princeton University Press, 2008.