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 \( T \) of trajectories of a nonlinear system (subject to a fixed initial condition). To work on the trajectory manifold, one projects curves \( \xi \) in the ambient Banach space onto the trajectory manifold, giving \( \eta = P(\xi) \in T \), by using a local linear time-varying trajectory tracking controller. Noting that the constrained and unconstrained optimization problems

\[
\min_{\xi \in T} h(\xi) \quad \text{and} \quad \min_{\xi} h(P(\xi))
\]

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 \( \tilde{h}(\xi) = h(P(\xi)) \). In particular, the local Newton update, valid in a neighborhood of a second order sufficient local minimum, is given by \( \xi_{i+1} = P(\xi_i + \zeta_i) \) where

\[
\zeta_i = \arg \min_{\zeta \in T_{\xi_i} T} \nabla^2 \tilde{h}(\xi_i) \cdot \zeta + \frac{1}{2} \nabla^2 \tilde{h}(\xi_i) \cdot (\zeta, \zeta)
\]

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

\[
\nabla^2 \tilde{h}(\xi) \cdot (\zeta, \zeta) = \nabla^2 h(\xi) \cdot (\zeta, \zeta) + \nabla h(\xi) \cdot \nabla^2 P(\xi) \cdot (\zeta, \zeta)
\]

(for \( \xi \in T, \zeta \in T_\xi T \)) 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/67072/2006), NAV-Control/FCT-PT (PTDC/EEA-ACR/65996/2006), Co3-AUVs (EU FP7 no. 231378), FCT-ISR/IST pluriannual 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 {a.saconc, ppedro}@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 throughout 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 \( x \). \( T_x M \) 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 \( \pi v_x = x \). 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 \( D_f : TM \to TN \) (or also as \( T f : TM \to TN \)). Tangent maps act naturally on tangent vectors. Given a vector \( v_x \in T_x M \), \( D_f(x) \cdot v_x \in T_{f(x)} N \) (or \( T_x f(v_x) \)) 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 \( D_f \cdot X : M \to TN \) (and \( T f(X) : M \to TN \) denotes at \( x \in M \) the tangent vector \( D_f(x) \cdot X(x) \in T_{f(x)} N \) (and \( T_x f(X(x)) \)). Given a diffeomorphism \( \varphi : M \to N \) the push-forward of a vector field \( X \) on \( M \) through \( \varphi \), denoted by \( \varphi_\ast X \), is the vector field on \( N \) defined by \( (\varphi_\ast X)(y) = T\varphi(X(\varphi^{-1}(y))) \), \( y \in N \). Given a diffeomorphism \( \varphi : M \to N \) the pull-back of a vector field \( Y \) on \( N \) through \( \varphi \), written as \( \varphi^\ast Y \), is the vector field on \( M \) defined by \( (\varphi^\ast Y)(x) = T\varphi^{-1}(Y(\varphi(x))) \), \( x \in M \), that is \( (\varphi^{-1})_\ast Y \). Given an affine connection \( \nabla \) on a manifold \( M \), we write \( \nabla_X Y \) 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 displacement along
a curve $\gamma(t), t \in I$, from $t = t_0$ to $t = t_1$ of a vector $V_0 \in T_x(\gamma(t_0))M$ is represented by $P^t_{t_0}V_0$. We also adopt (see Section VI) the notation $\nabla^2_YY \cdot X$ to mean $\nabla_XY$. Given a function $f : M \to N$, $D^2f(x) \cdot (v_x, w_x) \in T_f(x)N$ 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 $L_gx$ and $R_gx$, respectively. When convenient, we will adopt the shorthand notation $gz \equiv zg \equiv gv_x, vxg$ for, in the same order, $L_gx, R_gx, T_xL_g(v_x)$ and $T_xR_g(v_x)$. A left-invariant vector field on $G$ is a vector field such that $X(L_gx) = T_xL_g(X(x))$. Given a tangent vector at the identity $\dot{g} \in T_eG$, the symbol $X_{\dot{g}}$ means the left-invariant vector field defined by $X_{\dot{g}}(g) \equiv T_eL_g(\dot{g})$. The Lie algebra of $G$ is denoted by $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is identified with the tangent space $T_eG$ endowed with the Lie bracket operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, defined by $[g, h] \equiv [X_g, X_h](e)$, where the latter bracket is the Jacobi-Lie bracket of the left-invariant vector fields $X_g$ and $X_h$ 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 $Ad_g : \mathfrak{g} \to \mathfrak{g}$ and is the tangent map obtained differentiating $I_g(x)$ with respect to $x$ at $x = e$. We recall that $g = T_eG$. Furthermore, the adjoint representation of the Lie algebra $\mathfrak{g}$ onto itself is written as $ad_g : \mathfrak{g} \to \mathfrak{g}$ and it is obtained differentiating $Ad_g(\xi)$ with respect to $g$ at $g = e$, in the direction $\xi$. We recall that $ad_{g\xi} = [g, \xi]$. The exponential map is denoted by $\exp: \mathfrak{g} \to G$ and its inverse (in the 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 $\xi \in N_e \subseteq \mathfrak{g}$ the (right) trivialized tangent of $F$ at $\xi$ is the linear mapping $dF_\xi : \mathfrak{g} \to \mathfrak{g}$ defined by

$$dF_\xi := DF(\cdot) \cdot TRG_\eta,$$

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

$$dH_\xi := TR_{H(\xi)^{-1}}(DH(\cdot) \cdot \eta)$$

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 $u(t)$,

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

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

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

Remark. We would like to emphasize that the tangent bundle of a Lie group is itself a Lie group, with operation $(g_1, v_1), (g_2, v_2) \equiv (g_1g_2, g_1v_2 + g_2v_1) \in T_{g_1g_2}G$. 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 $e \in \mathbb{R}$ "small", consider the linear perturbation of the input defined as $u_c(t) \equiv u(t) + \varepsilon v(t)$. Indicating with $g_c$ the state trajectory associated with $u_c$, we have

$$\dot{g}_c(t) = g_c(t)\lambda(g_c(t), u_c(t), t),$$

$$g_c(0) = g_0.$$

In the (possibly small) interval $[0, T(\varepsilon)]$, the solution $g_c$ 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 $z_c(t), t \in [0, T(\varepsilon)]$, such that $g_c(t) = g(t)\exp(z_c(t)), t \in [0, T(\varepsilon)]$. The trajectory $z_c$ satisfies the following differential equation.

$$\dot{z}_c = d\log z_c(Ad_{z_c}\lambda(gx_c, u_c, t) - \lambda(g,u,t))$$

$$z_c(0) = 0.$$
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 $\alpha$ is an exponentially stable (state) trajectory of the closed loop system

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

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

$$u(t) = k(g(t), \alpha(t), \mu(t), t) = \mu(t) + K(t) \ln(g(t)-1)/\alpha(t)), \quad (15)$$
as a trajectory $\xi$ 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 $\xi$. It will be evident from next section that the linearization of (14) with feedback (15) around a state trajectory $\alpha$ 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 $P$): 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

$$\dot{g} = g\lambda_K(g, \xi(t), t), \quad (17)$$
$$\dot{u} = u_K(g, \xi(t), t), \quad (18)$$
$$g(0) = g_0, \quad (19)$$

where

$$\lambda_K(g, \xi, t) := \lambda(g, u_K(g, \xi, t)), \quad (20)$$
$$u_K(g, \xi, t) := \mu + K(t) \ln(g^{-1}/\alpha). \quad (21)$$

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

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 $g \times \mathbb{R}^m$. 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 $\xi$ in the direction $\zeta$, that is to study the mapping $P(\xi \exp(\epsilon \zeta))$, for $\epsilon \in \mathbb{R}$ “small”. From continuity of the mapping $P$, we can parameterize $P(\xi \exp(\epsilon \zeta))$ using the left-trivialized perturbed trajectory $\chi_\epsilon(t) \in g \times \mathbb{R}^m$, $t \geq 0$, defined by

$$P(\xi \exp(\epsilon \zeta)) = P(\xi) \exp(\epsilon \zeta). \quad (22)$$

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

$$N_\xi(\zeta) := \log(\exp(\epsilon \zeta)) \cdot \exp(\epsilon \zeta). \quad (23)$$

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

$$y_\epsilon = d \log_{y_0} [A \exp(u_0)]$$
$$\lambda_K(g \exp y_\epsilon, \xi \exp(\epsilon \zeta)) - \lambda_K(g, \xi, t)), \quad (24)$$
$$w_\epsilon(t) = u_K(g \exp y_\epsilon, \xi \exp(\epsilon \zeta)) - u_K(g, \xi, t), \quad (25)$$
$$y_\epsilon(0) = 0. \quad (26)$$

Proposition 5.2: The left-trivialized trajectory $\chi_\epsilon = N_\xi(\epsilon \zeta)$ can be expanded to first order as $\chi_\epsilon(t) = \epsilon \gamma(t) + R_2(\epsilon, t)$ with $R_2$ of order higher than one in $\epsilon$ and $\gamma(t) = (z(t), u(t))$, $t \geq 0$, satisfies

$$\gamma = D N_\xi(0) \cdot \zeta = \mathcal{P}(\xi)^{-1} \mathcal{D} \mathcal{P}(\xi) \cdot \zeta \quad (27)$$
and can be computed using

$$\dot{z} = A(\eta(t)) z + B(\eta(t)) v, \quad (28)$$
$$v = \nu + K \log(y_0 - 1) A \nu - 1 \beta - z, \quad (29)$$
$$z(0) = 0, \quad (30)$$
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 $\lambda$ 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 $P$.

Definition 6.1: An affine connection or covariant derivative $\nabla$ (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 $X$

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

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

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

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

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

$$D_t : T(\gamma) \rightarrow T(\gamma)$$

satisfying for each $V \in TM$,$ \nabla$ determines a unique operator

1) Linearity over $\mathbb{R}$

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

2) Product rule

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

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

$$D_tV(t) = (\nabla^\gamma_{t(t)}Y)(\gamma(t)).$$

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 $\nabla$. Given a curve $\gamma : I \rightarrow M$, the parallel displacement of the vector $V_0 \in T_{\gamma(t_0)}M, t_0 \in I$, along $\gamma$ at time $t_1 \in I$ is given by the unique vector field $V$ along $\gamma$, $V \in T(\gamma)$, that satisfies

$$D_tV(t) = 0, \quad t \in I,$$

$$V(t_0) = V_0,$$

(31)
evaluated at time $t_1$. We will denote such a (linear) transformation by $P_{\gamma}^{t_1-t_0}V_0 := V(t_1)$.

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

**VII. BI-IN-VARIANT 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 $G \times g$. Amongst all possible affine connections $\nabla$, left-invariant connections are those which commute with the push-forward of the left translation, that is

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

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 $\omega : g \times g \rightarrow g$ given by

$$\omega(g, \varsigma) = (\nabla_{x_\varsigma} x_\varsigma)(x), \ g, \varsigma \in g$$

(33)

for $X_{\omega}(g) := TL_g\omega$ (resp., := $TR_g\omega$).

The bilinear function $\omega$ appearing in (33) is termed the left (respectively, right) connection function for $\nabla$. 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 $\gamma(t) := \exp(tg_0)$ is a geodesic, that is, $\nabla_{\dot{\gamma}}\gamma(t) = 0$. For bi-invariant connections, the right and left-connection functions coincide and we have $\omega(Ad_g, \varsigma) = Ad_g\omega(g, \varsigma)$. See [11, Section 11].

Let $\gamma : I \rightarrow G$ 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

$$\omega(g, \varsigma) = 0,$$ 

$g, \varsigma \in g,$

(34)

$$P_{\gamma}^{t_1-t_0}V_0 = x_1x_0^{-1}V_0, \quad v_0 \in T_{x_0}G;$$

(35)

- The $(+)$ connection satisfies

$$\omega(g, \varsigma) = [g, \varsigma],$$ 

$g, \varsigma \in g,$

(36)

$$P_{\gamma}^{t_1-t_0}V_0 = v_0 x_0^{-1}x_1, \quad v_0 \in T_{x_0}G;$$

(37)

- The $(0)$ connection satisfies

$$\omega(g, \varsigma) = 1/2 [g, \varsigma],$$ 

$g, \varsigma \in g,$

(38)

$$P_{\gamma}^{t_1-t_0}V_0 = 1/2 (x_1x_0^{-1}V_0 + v_0x_0^{-1}x_1), \quad v_0 \in T_{x_0}G.$$ 

(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 $\nabla$ and $\nabla^\infty$, 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” $Df(\gamma(t)) \cdot V_1(t)$, for a curve $\gamma$ and a vector field $V_1$ along $\gamma_1$ in $M_1$.

Chosen $x \in M_1$ and two tangent vectors $v_x$ and $w_x \in T_xM_1$, let $\gamma_1 : I \to M_1$ be a smooth curve in $M_1$ such that $\gamma_1(t_0) = x$ and $\dot{\gamma}_1(t_0) = v_x$, $V_1$ a smooth vector field along $\gamma_1$ such that $V_1(t_0) = v_x$, and $V_2(t) := Df(\gamma(t_0)) \cdot V_1(t) \in T_{f(\gamma(t))}M_2$ a smooth vector field along the curve $\gamma_2(t) := f(\gamma(t))$ 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 direction $v_x$ and $w_x \in T_xM_1$ is the bilinear mapping $D^2f(x) : T_xM_1 \times T_xM_1 \to T_{f(x)}M_2$ defined as

$$D^2f(x) \cdot (v_x, w_x) = D_xv_x(0) - Df(\gamma(t_0)) \cdot D_xv_x(0),$$

(40)

where $D_xV_1$ and $D_xV_2$ denote the covariant differentiation with respect to $\nabla$ and $\nabla^\infty$, respectively.

**Corollary 8.1:** Denote by $\Pi$ and $\Pi^\infty$ the parallel displacements associated to $\nabla$ and $\nabla^\infty$, respectively. Then, equation (40) is equal (for $t = 0$) to

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( 2p^{t_{\gamma_2} + t + \epsilon} Df(\gamma(t + \epsilon)) \cdot 1_p^{t + \epsilon - t} V_1(t) \right.\left. - Df(\gamma(t)) \cdot V_1(t) \right),$$

(41)

**Proof:** The connection $\nabla^\infty$ allows us to compute the covariant derivative of the vector field $V_2$ along $\gamma_2$ as

$$(D_xV_2)(t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( 2p^{t_{\gamma_2} + t + \epsilon} V_2(t + \epsilon) - V_2(t) \right),$$

(42)

Equation (42) can be expanded into

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( 2p^{t_{\gamma_2} + t + \epsilon} Df(\gamma(t + \epsilon)) \cdot 1_p^{t + \epsilon - t} V_1(t) \right.\left. - Df(\gamma(t)) \cdot V_1(t) \right).$$

(43)

Adding and subtracting the term $2p^{t_{\gamma_2} + t + \epsilon} Df(\gamma(t + \epsilon)) \cdot 1_p^{t + \epsilon - t} V_1(t)$ inside the parenthesis of the previous expression, and noting that (in $TM_2$)

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( 2p^{t_{\gamma_2} - t + \epsilon} Df(\gamma(t + \epsilon)) \cdot 1_p^{t + \epsilon - t} V_1(t) \right.$$\n
$$\left. - Df(\gamma(t)) \cdot V_1(t) \right).$$

(43)

is equal to $Df(\gamma(t)) \cdot D_xV_1(t)$, 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 $f = id$ (the identity map), $D^2f$ reduces to the difference of the connections $\nabla$ and $\nabla^\infty$, that is $D^2f(x) \cdot (X(x), Y(x)) = (\nabla_X Y - \nabla_Y X)(x)$, which is known to be a $(1,2)$-tensor. 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 derivatives $D^3f$, $D^4f$, 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 $D^3 \log$ in computing the second order approximation of $N_\xi$ 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 $g \to \xi(g)$ a $g$-valued function. The following holds.

$$D^Xg \cdot \eta g = (D\xi(g) \cdot \eta g + 1/2 [\eta, \xi(g)])$$

$$D^Yg \cdot \eta g = (D\xi(g) \cdot \eta g + 1/2 [\xi(g), \eta])g.$$  

(44)

**Proposition 9.2:** Let $G \ni g \to \xi(g) \in \mathfrak{g}$ be defined as $g^{-1}X(g) = T_0L_g^{-1}X(g)$, with $X(g)$ a vector field. Then

$$D\xi(g) \cdot \eta g = g^{-1} \left[ D^Xg(\cdot) \cdot \eta g \right] + 1/2 [g^{-1}X(g), \eta].$$

**Proposition 9.3:** For each $g, \zeta \in \mathfrak{g}$, we have

$$D^2 \exp(0) \cdot (g, \zeta) = 0, \quad D^2 \log(e) \cdot (g, \zeta) = 0.$$  

**Proposition 9.4:** Let $t \to V(t)$ 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 $D_t W(t) = g D_t V(t)$.

X. QUADRATIC APPROXIMATION OF THE UNCONSTRAINED PROBLEM

Consider the cost functional $h(\xi) := \int_0^t l(\xi(\tau), \tau) \, d\tau$, where $l(\xi(t), \mu(t)) = \pi_1 \xi(t)$.

Using the Projection Operator $\Pi$, the functional $h$ over the space of curves in $G \times \mathbb{R}^m$ is constructed as

$$\hat{h}(\xi) = h(\Pi(\xi)).$$

(46)

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

$$h(\Pi(\exp \varepsilon \zeta)) = h(\Pi(\exp(N_\xi(\varepsilon \zeta)))),$$

(47)

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

\[
\begin{align*}
&h(P(\xi \exp(\varepsilon \zeta))) = h(P(\xi)) + \varepsilon Dh(P(\xi)) \cdot D_P(\xi) \cdot \xi \\
&+ 1/2 \varepsilon^2 \left[ D^2 h(P(\xi)) \cdot (D_P(\xi) \cdot \xi, D_P(\xi) \cdot \xi) \\
&+ Dh(P(\xi)) \cdot D^2 P(\xi) \cdot (\xi, \zeta) \right] + o(\varepsilon^2)
\end{align*}
\] (48)

where the first and second geometric derivative of \( h \) are

\[
D h(\xi) \cdot \xi = \int_0^t D_l(\xi(\tau), \tau) \cdot \xi(\tau) d\tau \\
+ D m(\pi_1 t_f) \cdot \pi_1(\xi(t_f))
\]

and

\[
D^2 h(\xi) \cdot (\xi_1, \xi_2) = \int_0^t D^2 l(\xi(\tau), \tau) \cdot (\xi(\tau) \xi_1(\tau), \xi(\tau) \xi_2(\tau)) d\tau \\
+ D^2 m(\pi_1 t_f) \cdot (T \pi_1(\xi(t_f)), T \pi_1(\xi(t_f)))
\]

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

\[
\begin{align*}
&h(P(\xi \exp(\varepsilon \zeta))) = h(P(\xi)) \\
&+ \varepsilon Dh(P(\xi)) \cdot P(\xi) D_N(\xi) \cdot \zeta \\
&+ 1/2 \varepsilon^2 \left[ D^2 h(P(\xi)) \cdot (P(\xi) D_N(\xi) \cdot \zeta, P(\xi) D_N(\xi) \cdot \zeta) \\
&+ Dh(P(\xi)) \cdot P(\xi) D^2 P(\xi) \cdot (D_N(\xi) \cdot \zeta, D_N(\xi) \cdot \zeta) \\
&+ D^2 N(\xi) \cdot (\zeta, \zeta) \right] + o(\varepsilon^2).
\end{align*}
\] (50)

From the definition of \( N_\xi \), we get

\[
D N(\xi) \cdot \zeta = P(\xi)^{-1} D P(\xi) \cdot \zeta
\] (52)

and, furthermore,

\[
D^2 N(\xi) \cdot (\zeta_1, \zeta_2) = P(\xi)^{-1} D^2 P(\xi) \cdot (\zeta_1, \zeta_2).
\] (53)

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

**Proposition 10.1:** The second geometric derivative of the \( P \) with respect to the (0) connection satisfies

\[
D^2 N(\xi) \cdot (\zeta_1, \zeta_2) = P(\xi)^{-1} D^2 P(\xi) \cdot (\zeta_1, \zeta_2).
\] (54)

Note that we write \( D^2 N \) instead of \( D^2 N_\xi \) to highlight the fact that \( N_\xi \) 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 \( D^2 \exp(0) \cdot (\varphi, \zeta) = -D^2 \log(e) \cdot (\varphi, \zeta) \), \( \forall \varphi, \zeta \in g \), for any choice of connection!

Summarizing the above discussion, we can conclude that the quadratic expansion of \( h \) is as follows.

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

XI. THE SECOND DERIVATIVE OF \( N_\xi \)

This section shows the relationship between the second derivative of the Local Projection Operator \( N \) and the second geometric derivative of the Projection Operator \( P \) at \( \xi \).

**Proposition 11.1:** Given a trajectory \( \eta = (g, u), \eta = P(\eta), \) the second derivative of the (left-trivialized) Local Projection Operator \( N_\xi \) around zero in the directions \( \zeta_1 \) and \( \zeta_2 \),

\[
(y, w) = D^2 N(y, u)(0) \cdot ((\beta_1, \nu_1), (\beta_2, \nu_2)) = D^2 N(0) \cdot (\zeta_1, \zeta_2) = P(\eta)^{-1} D^2 P(\eta) \cdot (\eta_1, \eta_2),
\]

is given by

\[
\begin{align*}
&\dot{y} = A(\eta)y + B(\eta)w - 1/2 \left( (\dot{\alpha}_1 \dot{\alpha}_2 + \dot{\alpha}_2 \dot{\alpha}_1) \lambda(\eta) \\
&- \dot{\alpha}_1 \dot{\alpha}_2 A(\eta) \dot{c}_2 + B(\eta) \nu_2) - \dot{\alpha}_2 A(\eta) \dot{c}_1 + B(\eta) \nu_1 \right) \\
&+ D^2 \lambda(\eta) \cdot (\eta_1, \eta_2),
\end{align*}
\] (55)

\[
w = -K(t) \left[ y + 1/2 \left( (\beta_1, \beta_2) + \dot{\beta}_2, \beta_1) \right) \right],
\]

with \( y(0) = 0, \dot{y}_i = (\dot{z}_1, \dot{v}_i) = D N(0) \cdot \zeta_i, i = \{1, 2\}, \) 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 were used to develop expressions for the basic objects needed for trajectory optimization.

We thank the anonymous reviewers for their valuable comments.

REFERENCES