

REDUCTION BY GROUP SYMMETRY OF SECOND ORDER VARIATIONAL PROBLEMS ON A SEMIDIRECT PRODUCT OF LIE GROUPS WITH POSITIVE DEFINITE RIEMANNIAN METRIC *

CLAUDIO ALTAFINI¹

Abstract. For a Riemannian structure on a semidirect product of Lie groups, the variational problems can be reduced using the group symmetry. Choosing the Levi-Civita connection of a positive definite metric tensor, instead of any of the canonical connections for the Lie group, simplifies the reduction of the variations but complicates the expression for the Lie algebra valued covariant derivatives. The origin of the discrepancy is in the semidirect product structure, which implies that the Riemannian exponential map and the Lie group exponential map do not coincide. The consequence is that the reduced equations look more complicated than the original ones. The main scope of this paper is to treat the reduction of second order variational problems (corresponding to geometric splines) on such semidirect products of Lie groups. Due to the semidirect structure, a number of extra terms appears in the reduction, terms that are calculated explicitly. The result is used to compute the necessary conditions of an optimal control problem for a simple mechanical control system having invariant Lagrangian equal to the kinetic energy corresponding to the metric tensor. As an example, the case of a rigid body on the Special Euclidean group is considered in detail.

Mathematics Subject Classification. 22F30, 70Q05, 93B29, 49J15, 70H30, 70H33.

Received March 27, 2003.

1. INTRODUCTION

The use of group symmetry to simplify the formulation of Euler-Lagrange equations defined on the tangent bundle of a Lie group G is well-known in the literature on geometric mechanics, see [19]. The reduction is based on factoring out the dependence from G in a G -invariant Lagrangian *i.e.*, in studying a variational problem on $\mathfrak{g} \simeq TG/G$ rather than on the whole of TG . Instead of Euler-Lagrange equations on TG , one obtains the Euler-Poincaré equations on $G \times \mathfrak{g}$. If the Lagrangian is given by kinetic energy only, then the Riemannian counterpart of this formulation corresponds to the reduction of the first variational formula. Assume that G is a semidirect product of a Lie group and a vector space, without nontrivial fixed points, and that the metric tensor \mathbb{I} is positive definite. Due to the semidirect product structure, such a metric cannot be biinvariant and

Keywords and phrases. Lie group, semidirect product, second order variational problems, reduction, group symmetry, geometric splines, optimal control.

* *Work done while the author was with the Division of Optimization and Systems Theory, Royal Institute of Technology, Stockholm, Sweden. Support from the Swedish Foundation for Strategic Research through the Center for Autonomous Systems at KTH is acknowledged.*

¹ SISSA-ISAS, International School for Advanced Studies, via Beirut 2-4, 34014 Trieste, Italy; e-mail: altafini@sissa.it

© EDP Sciences, SMAI 2004

therefore the Levi-Civita connection induced by \mathbb{I} is (in the language of [14], Chap. X) neither natural nor canonical. In this case, in fact, the natural connection is pseudo-Riemannian *i.e.*, the corresponding quadratic form has to have both positive and negative eigenvalues.

The advantage of choosing \mathbb{I} positive definite (beside being compatible with simple mechanical systems having G as configuration space) is that the reduction of the variations of curves can be carried out quite easily. In fact, for families of proper variations the symmetry lemma, expressing the commutativity of the variational fields along the main and transverse curves, still holds after the reduction since all the vector fields involved admit invariant expressions. What gets more complicated is the reduction of the covariant derivatives, as the notion of parallel transport given by the Riemannian connection does not fit with the reduction process. This is due to the difference between the Riemannian exponential map associated with \mathbb{I} and the Lie group exponential map, and to the consequent mismatch between the two types of one-parameter subgroups. So, for example, geodesics of \mathbb{I} do not correspond to one-parameter subgroups of G . In spite of this complication, the reduction of the first order variational formula (*i.e.*, the Euler-Poincaré equations) is still quite easy to obtain and its advantage in practical applications over the full Euler-Lagrange equations well-documented (for their exploitation in Robotic applications see [4, 5]). The scope of this paper is to treat in a similar way the reduction of second order variational problems on G that can be associated with \mathbb{I} .

It is an elementary fact in calculus of variations that extremals of the energy functional give geodesic motion through the first variational formula. This leads to Euler-Lagrange equations or to Euler-Poincaré after the reduction. The corresponding necessary conditions for a cost function which is the L_2 norm of the acceleration were obtained in [10, 22, 24] for Riemannian manifolds and compact semisimple Lie groups. They resemble the equations for the Jacobi fields associated with the connection and they generalize to Riemannian manifolds the standard procedures to generate cubic splines in \mathbb{R}^n .

The reduction process can be seen as the projection map $\pi : TG \rightarrow \mathfrak{g}$ of a globally trivial principal fiber bundle (induced by left or right invariance of G) with base manifold \mathfrak{g} and structure group G . The mismatch between Lie group exponential map and Riemannian exponential map implies that the horizontal vectors determined by the Levi-Civita connection on TG are not anymore horizontal in the fiber bundle (*i.e.*, they do not reduce “exactly” as in Lie groups with biinvariant metric). The component which becomes vertical after the reduction belongs to the vector space (in the semidirect decomposition of G) and gives an extra drift term to the Euler-Poincaré equations with respect to the full Euler-Lagrange equations. These problems due to the mismatch of the two exponential maps are amplified when computing the reduction of the second order variational problem and several extra terms appear, which we compute explicitly.

As an application of the result, we consider a closely related optimal control problem for a (fully actuated) simple mechanical control system [16]. The problem we consider here is the so-called “ C^2 dynamical interpolation problem” of [6, 10] in which the cost functional is the same “minimum acceleration” of the second order variational problem. If the actuators are body fixed, then they form a left-invariant codistribution in the cotangent bundle which fits in with (and motivates further) the reduction procedure. As a matter of fact, one of the long term motivations behind this line of research is the generation of smooth trajectories for (fully actuated) mechanical control systems composed of kinetic energy alone and that can be modeled as actuated rigid bodies evolving on the Special Euclidean group $SE(3)$, see [5, 21, 25, 26] for a general overview and [1] for a particular application to Robotic Manipulation. For the particular case of $SE(3)$, we carry out all the calculations in detail.

The paper is organized as follows. Sections 2–5 contain known results presented in a uniform way: in Sections 2, 4 and 5, we collect the background material on Riemannian geometry, on Riemannian geometry on Lie groups and on the reduction procedure for the Hamilton principle respectively, while in Section 3 the second order variational problem is formulated. Sections 6–8 contain the original contribution of the paper: the reduction of the second order variational problem in Section 6; its application as an optimal control problem for a mechanical control system in Section 7 and its explicit calculation for the $SE(3)$ case in Section 8.

2. PRELIMINARIES OF RIEMANNIAN GEOMETRY

The material in this section can be found in any book on Riemannian geometry, for example [9, 11, 15]. A *Riemannian metric* on a smooth n -dimensional manifold \mathcal{M} is a 2-tensor field \mathbb{I} that is symmetric and positive definite. \mathbb{I} determines an inner product $\langle \cdot, \cdot \rangle$ on each tangent space $T_x\mathcal{M}$, $x \in \mathcal{M}$. Call $\mathcal{D}(\mathcal{M})$ the space of smooth sections of $T\mathcal{M}$ whose elements are smooth vector fields on \mathcal{M} . An *affine connection* ∇ is a map taking each pair of vector fields X and Y to another vector field $\nabla_X Y$, called *covariant derivative* of Y along X , such that for $f \in C^\infty(\mathcal{M})$

- (1) $\nabla_X Y$ is \mathbb{R} -bilinear in X and Y ;
- (2) $\nabla_{fX} Y = f\nabla_X Y$;
- (3) $\nabla_X (fY) = f\nabla_X Y + (\mathcal{L}_X f)Y$

where $\mathcal{L}_X f$ is the Lie derivative of f along X .

Given a curve $\gamma(t)$ and a vector field X , the covariant derivative of X along γ is $\frac{DX}{dt} = \nabla_{\dot{\gamma}(t)} X$. In coordinates $x = (x^1, \dots, x^n)$, if $X_x = X^i \frac{\partial}{\partial x^i}$, $Y_x = Y^i \frac{\partial}{\partial x^i} \in T_x\mathcal{M}$, the covariant derivative is

$$(\nabla_X Y)^k = \frac{\partial Y^k}{\partial x^i} X^i + \Gamma_{ij}^k X^i Y^j \tag{1}$$

where $(\nabla_X Y)_x = (\nabla_X Y)^k \frac{\partial}{\partial x^k}$ and the n^3 quantities Γ_{ij}^k are called *Christoffel symbols* and are given (at x) by $\nabla_{\frac{\partial}{\partial x^i}} (\frac{\partial}{\partial x^j}) = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$. For a generic smooth curve $\gamma(t) \in \mathcal{M}$ the quantity $\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = \frac{D}{dt} \left(\frac{d\gamma}{dt} \right)$ represents the acceleration. The length of the smooth curve γ is measured by the functional

$$\ell(\gamma) = \int_{t_0}^{t_f} \langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle^{\frac{1}{2}} dt. \tag{2}$$

A vector field Y is said *parallel transported* along γ if $\frac{DY}{dt} = 0$. In particular, if $\dot{\gamma}$ is parallel along γ , then γ is called a *geodesic*:

$$\frac{D}{dt} \left(\frac{d\gamma}{dt} \right) = \nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0. \tag{3}$$

Geodesic motion corresponds to constant velocity and it gives an extremum of the length functional (2), as well as of the kinetic energy integral $\int_{t_0}^{t_f} \langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle dt$. The condition for parallel transport of the vector Y along γ in coordinates becomes $\frac{dY^k}{dt} + \Gamma_{ij}^k \dot{x}^i Y^j = 0$ and the one for geodesic motion

$$\ddot{x}^k + \Gamma_{ij}^k \dot{x}^i \dot{x}^j = 0. \tag{4}$$

As \mathbb{I} is symmetric positive definite, it is compatible with the kinetic energy of a so-called simple mechanical system [16] for which the geodesic equations (3) are normally called Euler-Lagrange equations.

The fundamental theorem of Riemannian geometry says that given \mathbb{I} on a manifold \mathcal{M} there exists a unique affine connection ∇ on \mathcal{M} such that

- (1) ∇ is torsion free:

$$\nabla_X Y - \nabla_Y X = [X, Y]; \tag{5}$$

- (2) the parallel transport is an isometry

$$\mathcal{L}_Z \langle X, Y \rangle = \langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle \tag{6}$$

for all $X, Y, Z \in \mathcal{D}(\mathcal{M})$. Such a connection is called the *Levi-Civita connection*. From (6), we get the Koszul formula:

$$\begin{aligned} \langle Z, \nabla_X Y \rangle = & \frac{1}{2} (\mathcal{L}_Y \langle X, Z \rangle + \mathcal{L}_X \langle Z, Y \rangle - \mathcal{L}_Z \langle X, Y \rangle - \langle [Y, Z], X \rangle + \\ & + \langle [Z, X], Y \rangle + \langle [X, Y], Z \rangle). \end{aligned} \tag{7}$$

A vector field X on a Riemannian manifold is called a *Killing vector field* (or an *infinitesimal isometry*) if the local one-parameter subgroup of transformations generated by X via the exponential map of the connection in a neighborhood of each point consists of isometries. We have the following equivalent characterizations:

Proposition 2.1 ([14], Prop. 3.2, Chap. VI). *Given a vector field X on a Riemannian manifold with metric connection $(\mathcal{M}, \mathbb{I})$, the following are equivalent:*

- (1) X is a Killing vector field;
- (2) $\mathcal{L}_X \mathbb{I} = 0$;
- (3) the Killing equation holds:

$$\mathcal{L}_X \langle Y, Z \rangle = \langle [X, Y], Z \rangle + \langle Y, [X, Z] \rangle \quad \forall Y, Z \in \mathcal{D}(\mathcal{M}) \tag{8}$$

or equivalently

$$\langle \nabla_Y X, Z \rangle + \langle Y, \nabla_Z X \rangle = 0 \quad \forall Y, Z \in \mathcal{D}(\mathcal{M}). \tag{9}$$

- (4) the linear map A_X given by $A_X Y = -\nabla_Y X$, $Y \in \mathcal{D}(\mathcal{M})$ is skew symmetric with respect to $\langle \cdot, \cdot \rangle$ everywhere in \mathcal{M} , i.e., $\langle A_X Y, Z \rangle + \langle Y, A_X Z \rangle = 0 \quad \forall Y, Z \in \mathcal{D}(\mathcal{M})$.

The “measure” of the failure of the second covariant derivative to commute is expressed geometrically by the notion of *curvature*, i.e., the map $R : \mathcal{D}(\mathcal{M}) \times \mathcal{D}(\mathcal{M}) \times \mathcal{D}(\mathcal{M}) \rightarrow \mathcal{D}(\mathcal{M})$ defined by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z. \tag{10}$$

In coordinates, the coefficient $R_{ijk}{}^l$ of $R \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) \frac{\partial}{\partial x^k} = R_{ijk}{}^l \frac{\partial}{\partial x^l}$ are given by

$$R_{ijk}{}^l = \left(\frac{\partial \Gamma_{jk}^l}{\partial x^i} - \frac{\partial \Gamma_{ik}^l}{\partial x^j} \right) + (\Gamma_{jk}^m \Gamma_{im}^l - \Gamma_{ik}^m \Gamma_{jm}^l).$$

2.1. The variational principle of Hamilton

The geodesic equation (3) can be obtained from standard calculus of variation on the Riemannian manifold $(\mathcal{M}, \mathbb{I})$, see for example [15]. Given the curve $\gamma : [t_0, t_f] \rightarrow \mathcal{M}$, consider *proper variations* of γ i.e., the family of fixed end-point curves $\mathcal{G} : (-\epsilon, \epsilon) \times [t_0, t_f] \rightarrow \mathcal{M}$ such that

$$\mathcal{G}_0(t) = \mathcal{G}(s, t)|_{s=0} = \gamma(t) \quad \forall t \in [t_0, t_f]$$

and

$$\mathcal{G}_s(t_0) = \mathcal{G}(s, t_0)|_{s=\text{const}} = \gamma(t_0), \quad \mathcal{G}_s(t_f) = \mathcal{G}(s, t_f)|_{s=\text{const}} = \gamma(t_f) \quad \forall s \in (-\epsilon, \epsilon).$$

In the family of curves \mathcal{G} , the curves with fixed s , $\mathcal{G}_s(t) = \mathcal{G}(s, t)|_{s=\text{const}}$, are called *main curves* and those with fixed t , $\mathcal{G}^{(t)}(s) = \mathcal{G}(s, t)|_{t=\text{const}}$, *transverse curves*. At infinitesimal level, we call a *variation field* $\delta\gamma$ the tangent vector with respect to a transverse variation taken for a fixed $t \in [t_0, t_f]$ and computed at $s = 0$:

$$\delta\gamma(t) = \left. \frac{d}{ds} \mathcal{G}^{(t)}(s) \right|_{s=0}.$$

The variation field is proper if $\delta\gamma(t_0) = \delta\gamma(t_f) = 0$. If \mathcal{G} is proper, then $\delta\gamma$ is also proper. It is a standard result that any C^2 vector field along γ is the variation field of some variation of γ , and that if $\delta\gamma$ is proper so is the corresponding variation. This is proven *via* the Riemannian exponential map Exp associated with the Levi-Civita connection ∇ : the variation corresponding to a vector field $V(t)$ along $\gamma(t)$ will be of the type $\mathcal{G}(s, t) = \text{Exp}(sV(t))$. In fact, for a fixed $\bar{t} \in [t_0, t_f]$, if we have $\mathcal{G}^{(\bar{t})}(s) = \text{Exp}(sV(\bar{t}))$ then

$$\delta\gamma(\bar{t}) = \left. \frac{d}{ds} \text{Exp}(sV(\bar{t})) \right|_{s=0} = V(\bar{t}).$$

The variations $\mathcal{G}(, s)$ are continuous in $(-\epsilon, \epsilon) \times [t_0, t_f]$ and smooth on rectangles $(-\epsilon, \epsilon) \times [t_j, t_{j+1}]$ for some subdivision $t_0 \leq t_1 \leq \dots \leq t_k \leq t_f$. Another standard result is the *symmetry lemma*, that allows one to exchange the order of the mixed second order derivatives along main and transverse curves. Calling

$$S(s, t) = \frac{d}{ds} \mathcal{G}^{(t)}(s) \quad \text{and} \quad T(s, t) = \frac{d}{dt} \mathcal{G}_s(t) \tag{11}$$

(so that $S(0, t) = \delta\gamma(t)$ and $T(0, t) = \dot{\gamma}(t)$), we have $\nabla_S T = \nabla_T S$. For a torsion-free connection, this implies, from (5), that the vector fields T and S commute¹ $[T, S] = 0$. Furthermore, from (6) we have

$$\frac{d}{dt} \langle S, T \rangle = \mathcal{L}_T \langle S, T \rangle = \langle \nabla_T S, T \rangle + \langle S, \nabla_T T \rangle. \tag{12}$$

The Hamilton principle for the functional $\ell(\gamma)$ gives the curve $\gamma(t)$ for which ℓ is stationary under proper variations. Considering, for the sake of simplicity, in place of ℓ the energy functional

$$\mathcal{E}(\gamma) = \frac{1}{2} \int_{t_0}^{t_f} \langle \dot{\gamma}, \dot{\gamma} \rangle dt$$

we have

$$\begin{aligned} \left. \frac{d}{ds} \mathcal{E}(\mathcal{G}_s(t)) \right|_{s=0} &= \left. \frac{1}{2} \frac{d}{ds} \int_{t_0}^{t_f} \langle T(s, t), T(s, t) \rangle dt \right|_{s=0} \\ &= \left. \int_{t_0}^{t_f} \langle \nabla_S T, T \rangle dt \right|_{s=0} \\ &= \int_{t_0}^{t_f} \langle \nabla_{\delta\gamma} \dot{\gamma}, \dot{\gamma} \rangle dt = \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}} \delta\gamma, \dot{\gamma} \rangle dt && \text{by the symmetry lemma} \\ &= \int_{t_0}^{t_f} \left(\frac{d}{dt} \langle \delta\gamma, \dot{\gamma} \rangle - \langle \nabla_{\dot{\gamma}} \dot{\gamma}, \delta\gamma \rangle \right) dt && \text{by (12)} \\ &= \langle \delta\gamma, \dot{\gamma} \rangle \Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}} \dot{\gamma}, \delta\gamma \rangle dt. \end{aligned}$$

Since $\delta\gamma(t_0) = \delta\gamma(t_f) = 0$, we obtain the first variation formula

$$\left. \frac{d}{ds} \mathcal{E}(\mathcal{G}_s(t)) \right|_{s=0} = 0 \iff \nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0 \tag{13}$$

which corresponds to the Euler-Lagrange equations for a Lagrangian equal to kinetic energy only.

¹Here and in the following: when needed, operations involving T and S will be intended on rectangles of the type $(-\epsilon, \epsilon) \times [t_j, t_{j+1}]$.

Considering only variations through geodesics, *i.e.*, families $\mathcal{G}(s, t)$ such that all the main curves $\mathcal{G}_s(t)$ are geodesics, a *Jacobi field* V is a vector field along γ satisfying the *Jacobi equation*

$$\nabla_{\dot{\gamma}}^2 V + R(V, \dot{\gamma})\dot{\gamma} = 0. \tag{14}$$

A vector field is a Jacobi field if and only if it is the variation field of some variations of γ among geodesics. The Jacobi equation is essentially a linear system of second order differential equations in V along γ . If properly initialized (its initial values being $\gamma(t_0)$, $V(t_0)$ and $\nabla_{\dot{\gamma}} V(t_0)$), then in the “variations through geodesics” case it has a unique solution for all t .

2.2. Second order structures on a Riemannian manifold

Assume that the coordinate chart x^1, \dots, x^n is valid in a neighborhood U of $x \in \mathcal{M}$. If $v \in T_x \mathcal{M}$ is a tangent vector, its coordinates description is naturally given by $v = v^i \frac{\partial}{\partial x^i}$. If $\tau : T\mathcal{M} \rightarrow \mathcal{M}$ is the tangent bundle projection, $(x^1, \dots, x^n, v^1, \dots, v^n)$ are called induced coordinates on $\tau^{-1}(U)$ and they provide a basis of tangent vectors of $T_{(x,v)}T\mathcal{M}$: $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}, \frac{\partial}{\partial v^1}, \dots, \frac{\partial}{\partial v^n})$. By taking the tangent map τ_* of the projection τ at the point (x, v) of $T\mathcal{M}$, $\tau_* : T_{(x,v)}T\mathcal{M} \rightarrow T_{x=\tau(v)}\mathcal{M} = \tau^{-1}(x)$, one can define the *vertical subspace* of the tangent bundle at (x, v)

$$\mathcal{V}_{(x,v)} = \ker \tau_* = \{w \in T_{(x,v)}T\mathcal{M} \text{ s.t. } \tau_*(w) = 0 \in T_x \mathcal{M}\}.$$

The vertical subspace is the subspace of $T_{(x,v)}T\mathcal{M}$ whose vectors are tangent to the fiber $\tau^{-1}(x) = T_x \mathcal{M}$. Such vectors are called *vertical lifts* and can be computed as follows: given the tangent vector $u \in T_x \mathcal{M}$ the vertical lift u^v of u from $T_{\tau(v)}\mathcal{M}$ to $T_{(x,v)}T\mathcal{M}$ is

$$u^v = \left. \frac{d}{dt} (v + tu) \right|_{t=0}.$$

For example $(\frac{\partial}{\partial x^i})^v = (\frac{\partial}{\partial v^i})$.

The complementary subspace to $\mathcal{V}_{(x,v)}$ in $T_{(x,v)}T\mathcal{M}$, in order to be identified, requires a notion of parallelism to be defined, for example through the Levi-Civita connection ∇ . The *horizontal lift* of $u \in T_x \mathcal{M}$ to a tangent vector on $T_{(x,v)}T\mathcal{M}$, in fact, is defined *via* the parallel transport of a vector field $V \in \mathcal{D}(\mathcal{M})$ such that $V(0) = v$ along a curve $\sigma(t) \in \mathcal{M}$ such that $\sigma(0) = x$ and $\dot{\sigma}(0) = u$ (see [9], Chap. 13). In fact, calling $\sigma^h = (\sigma, V)$ the horizontal lift of the curve σ through (x, v) , the condition $\nabla_u V|_{t=0} = 0$ (in coordinates $\dot{V}^i + \Gamma_{jk}^i V^j u^k = 0$) provides an expression for the derivative of V at $t = 0$ and the horizontal lift u^h of u from $T_x \mathcal{M}$ to $T_{(x,v)}T\mathcal{M}$ can be defined as the tangent vector to σ^h at $t = 0$:

$$\frac{d\sigma^h}{dt} = u^h \quad \text{s.t.} \quad \sigma^h(0) = (\sigma(0), V(0)).$$

If $u = u^i \frac{\partial}{\partial x^i}$, its expression in coordinates

$$\begin{aligned} \dot{\sigma} &= u \\ \dot{v}^i &= -\Gamma_{jk}^i v^j u^k \end{aligned}$$

or

$$u^h = u^k \frac{\partial}{\partial x^k} - \Gamma_{jk}^i v^j u^k \frac{\partial}{\partial v^i}. \tag{15}$$

Since $\tau_*(u^h) = u$, horizontal lifts are complementary to the vertical subspace and form the *horizontal subspace* $\mathcal{H}_{(x,v)}$ of $T_{(x,v)}T\mathcal{M}$. Equivalently, $\mathcal{H}_{(x,v)}$ can be defined in terms of sections of the tangent bundle, *i.e.*, of

smooth maps $\varsigma : \mathcal{M} \rightarrow T\mathcal{M}$ such that $\tau(\varsigma(x)) = x \ \forall x \in \mathcal{M}$. This is done by taking the push forward at $u \in T_{\tau(v)}\mathcal{M}$ of the sections that are parallel transported along the vector u

$$\mathcal{H}_{(x,v)} = \{\overline{\varsigma_* u} \text{ s.t. } \nabla_u \varsigma = 0\}. \tag{16}$$

For a curve $\gamma \in \mathcal{M}$ which is a geodesic, the horizontal lift $(\gamma, \dot{\gamma})$ is also called the natural lift. In this case u coincides with v and therefore the tangent vector to $(\gamma, \dot{\gamma})$ at $(x = \gamma(0), v = \dot{\gamma}(0))$ is the horizontal lift $v^h \in T_{(x,v)}T\mathcal{M}$ of v . The vector field Γ on $T\mathcal{M}$ such that $\Gamma_{(x,v)} = v^h$ is called the *geodesic spray* of the connection. From (15), using the same coordinate notation as above for $\gamma(t)$

$$\Gamma_{(x,v)} = v^k \frac{\partial}{\partial x^k} - \Gamma_{jk}^i v^j v^k \frac{\partial}{\partial v^i}. \tag{17}$$

Γ is characterized by integral curves that are natural lifts of geodesics. Written as a system of first order equations, the integral curves of Γ are (compare with (4))

$$\begin{aligned} \dot{x}^k &= v^k \\ \dot{v}^k &= -\Gamma_{ij}^k v^i v^j. \end{aligned}$$

3. A SECOND ORDER VARIATIONAL PROBLEM

Following [6, 22, 26], the problem of constructing trajectories between given initial and final position and velocity data on \mathcal{M} can be formulated as an optimization problem on a Riemannian manifold, taking as cost functional the square of the L^2 norm of the acceleration:

$$\mathcal{J} = \frac{1}{2} \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle dt. \tag{18}$$

\mathcal{J} has extremals that are generalizations to Riemannian manifolds of Euclidean cubic splines. Its first variation gives the necessary conditions for curves to be extremals.

Theorem 3.1 [10, 22]. *A necessary condition for a smooth curve $\gamma(t) \in \mathcal{M}$, $t \in [t_0, t_f]$, such that $\gamma(t_0) = g_0$, $\gamma(t_f) = g_f$, $\left. \frac{d\gamma}{dt} \right|_{t=t_0} = v_0$ and $\left. \frac{d\gamma}{dt} \right|_{t=t_f} = v_f$, to be an extremum of \mathcal{J} is that*

$$\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \dot{\gamma} + R(\nabla_{\dot{\gamma}} \dot{\gamma}, \dot{\gamma}) \dot{\gamma} = 0. \tag{19}$$

Proof. The proof has already appeared in the above mentioned references (see also [26]). It is repeated here only for the sake of completeness. It follows the same arguments used in finding the critical curves of the energy functional. Furthermore, it makes use of the following symmetry of the curvature tensor:

$$\langle R(V, W)Z, U \rangle = \langle R(U, Z)W, V \rangle \tag{20}$$

$$\begin{aligned}
 \frac{d}{ds} \mathcal{J}(\mathcal{G}_s(t)) \Big|_{s=0} &= \int_{t_0}^{t_f} \langle \nabla_S \nabla_T T, \nabla_T T \rangle dt \Big|_{s=0} \\
 &= \int_{t_0}^{t_f} \langle \nabla_T \nabla_S T + R(S, T)T, \nabla_T T \rangle dt \Big|_{s=0} && \text{by (10)} \\
 &= \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}} \nabla_{\delta\gamma} \dot{\gamma} + R(\delta\gamma, \dot{\gamma})\dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle dt \\
 &= \int_{t_0}^{t_f} \left(\frac{d}{dt} \langle \nabla_{\delta\gamma} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle - \langle \nabla_{\dot{\gamma}}^2 \dot{\gamma}, \nabla_{\delta\gamma} \dot{\gamma} \rangle + \langle R(\delta\gamma, \dot{\gamma})\dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle \right) dt && \text{by (12)} \\
 &= \langle \nabla_{\delta\gamma} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle \Big|_{t_0}^{t_f} + \\
 &\quad + \int_{t_0}^{t_f} \left(-\frac{d}{dt} \langle \nabla_{\dot{\gamma}}^2 \dot{\gamma}, \delta\gamma \rangle + \langle \nabla_{\dot{\gamma}}^3 \dot{\gamma}, \delta\gamma \rangle + \langle R(\delta\gamma, \dot{\gamma})\dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle \right) dt && \text{by (12)} \\
 &= (\langle \nabla_{\delta\gamma} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle - \langle \nabla_{\dot{\gamma}}^2 \dot{\gamma}, \delta\gamma \rangle) \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} (\langle \nabla_{\dot{\gamma}}^3 \dot{\gamma}, \delta\gamma \rangle + \langle R(\nabla_{\dot{\gamma}} \dot{\gamma}, \dot{\gamma})\dot{\gamma}, \delta\gamma \rangle) dt. && \text{by (20)}
 \end{aligned}$$

Since the variation is assumed proper, $\delta\gamma$ and $\nabla_{\delta\gamma} \dot{\gamma}$ both vanish at the end points and the result follows. \square

4. RIEMANNIAN CONNECTION ON A SEMIDIRECT PRODUCT OF LIE GROUPS

Consider a Lie group G and its left translation action $L_g, g \in G$. A vector field X on G is called left invariant if for every $g \in G$ we have $(T_h L_g) X_h = X_{gh} \forall h \in G$. The set $\mathfrak{X}_L(G)$ of left invariant vectors on G is isomorphic to the Lie algebra $\mathfrak{g} = T_e G$. Call \mathfrak{X} the Lie algebra evaluated vector: $\mathfrak{X} = X_e \in \mathfrak{g}$.

The Lie group is made in a Riemannian manifold by defining an inner product on $T_e G = \mathfrak{g}$ and propagating it on TG by left translation. This makes G automatically into a complete homogeneous Riemannian manifold. As mentioned above, we are interested in the case in which the metric tensor \mathbb{I} is compatible with the kinematic energy of a simple mechanical system having G as configuration space, and therefore we restrict to symmetric positive definite \mathbb{I} . In this case, the geodesics of the Levi-Civita connection are the solutions of the Euler-Lagrange equations for an invariant Lagrangian function corresponding to the kinetic energy.

The class of Lie groups we consider here has the structure of a semidirect product of a Lie group K and a vector space $V: G = K \ltimes V$, see [13]. As a manifold, G is the Cartesian product of K and V , but the Lie group multiplication includes the linear action of K on $V, K \rightarrow \text{Aut}(V)$, so that the group multiplication looks like

$$(k_1, u_1)(k_2, u_2) = (k_1 k_2, u_1 + k_1 u_2) \quad k_1, k_2 \in K, \quad u_1, u_2 \in V.$$

Consequently, the Lie algebra \mathfrak{g} of G includes the induced action $\mathfrak{k} \rightarrow \text{End}(V)$ and is therefore the semidirect sum of \mathfrak{k} and V with Lie bracket

$$[[\mathcal{K}_1, v_1], [\mathcal{K}_2, v_2]] = [[\mathcal{K}_1, \mathcal{K}_2], \mathcal{K}_1 v_2 - \mathcal{K}_2 v_1] \quad \mathcal{K}_1, \mathcal{K}_2 \in \mathfrak{k}, \quad v_1, v_2 \in V.$$

Obviously, V forms an ideal in \mathfrak{g} since it is abelian and therefore $[\mathcal{K}, v] \in V$ for all $\mathcal{K} \in \mathfrak{k}$ and $v \in V$.

We assume that V has no nonzero fixed points under K and that \mathfrak{k} has no nontrivial ideals.

Since K acts linearly on the vector space V , the whole Lie group G acts affinely on V :

$$(k, u_1)u_2 = k u_2 + u_1 \quad \forall (k, u_1) \in G \quad u_2 \in V.$$

Hence, if $\mathcal{Y} = (\mathcal{K}, v) \in \mathfrak{g}$ and $u \in V$, the infinitesimal generator of the one-parameter subgroup on $V, \phi_{\mathcal{Y}}(t)u = e^{t\mathcal{Y}}u$, is the affine vector field $\mathcal{Y}_G(u) = \mathcal{K}u + v$.

The Levi-Civita connection ∇ , being defined from a left-invariant metric, retains the left-invariant property along the coordinate directions of an invariant basis on TG . Calling $A_i \in \mathfrak{g}$ the elements of an orthonormal

basis of left invariant vector fields:

$$\nabla_{gA_i} (gA_j) = g\nabla_{A_i} A_j = \Gamma_{ij}^k gA_k \tag{21}$$

for all $g \in G$. From (1), since the Γ_{ij}^k are not tensorial, left invariance of the connection has to be intended with respect to affine transformations, *i.e.*, if X is an infinitesimal affine transformation and ϕ_X the corresponding local one-parameter group of local transformations in G generated by X (see Prop. 1.4, Chap. VI of [14]): $(\phi_X)_* (\nabla_Y Z) = \nabla_{(\phi_X)_*(Y)} (\phi_X)_*(Z) \forall Y, Z \in \mathcal{D}(G)$. The infinitesimal generator above $\mathcal{Y}_G(u)$ is an example of how one-parameter subgroups affinely generated emerge in a semidirect product.

A crucial point in the reduction program which will be carried out below is the understanding of the relation between the operations of Lie and covariant derivative and the left invariance of the vector fields involved. For example, in correspondence of $X \in \mathcal{D}(G)$ and $\mathcal{Y}, \mathcal{Z} \in \mathfrak{g}$, the equation (6) for the parallel transport simplifies to

$$\mathcal{L}_X \langle \mathcal{Y}, \mathcal{Z} \rangle = \langle \nabla_X \mathcal{Y}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \nabla_X \mathcal{Z} \rangle = 0 \tag{22}$$

since $\langle \mathcal{Y}, \mathcal{Z} \rangle$ is a constant. Hence if X is Killing the equation (8) becomes:

$$\mathcal{L}_X \langle \mathcal{Y}, \mathcal{Z} \rangle = \langle [X, \mathcal{Y}], \mathcal{Z} \rangle + \langle \mathcal{Y}, [X, \mathcal{Z}] \rangle = 0. \tag{23}$$

A constant metric quadratic form like \mathbb{I} , interpreted as an inertia tensor, is a map $\mathbb{I} : \mathfrak{g} \rightarrow \mathfrak{g}^*$ the dual of \mathfrak{g} . Using ad_X^* , the dual of ad_X , defined as $(\text{ad}_X \mathcal{Z}; \psi) = (\mathcal{Z}; \text{ad}_X^* \psi)$, $\mathcal{X}, \mathcal{Z} \in \mathfrak{g}$, $\psi \in \mathfrak{g}^*$ and $(\cdot; \cdot)$ indicating the \mathbb{R} -valued standard pairing between a Lie algebra and its dual, we get

$$\langle \text{ad}_X \mathcal{Z}, \mathcal{Y} \rangle = (\text{ad}_X \mathcal{Z}; \mathbb{I} \mathcal{Y}) = (\mathcal{Z}; \text{ad}_X^* \mathbb{I} \mathcal{Y}) = \langle \mathcal{Z}, \mathbb{I}^{-1} \text{ad}_X^* \mathbb{I} \mathcal{Y} \rangle. \tag{24}$$

We need to adapt to the case of trivially reductive homogeneous spaces given by the left action of a Lie group on itself the Theorem 3.3, Chapter X of [14]:

Theorem 4.1. *Given (G, \mathbb{I}) , the Riemannian connection for \mathbb{I} is expressed as*

$$\nabla_X \mathcal{Y} = \frac{1}{2} [X, \mathcal{Y}] + U(X, \mathcal{Y}) \tag{25}$$

where $U(X, \mathcal{Y})$ is the symmetric bilinear mapping $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ defined by

$$\langle U(X, \mathcal{Y}), \mathcal{Z} \rangle = \frac{1}{2} (\langle \text{ad}_Z X, \mathcal{Y} \rangle + \langle X, \text{ad}_Z \mathcal{Y} \rangle) \tag{26}$$

for all $X, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}$.

Applying the pairing (24) to $U(X, \mathcal{Y})$ then we have:

Proposition 4.2. *The left-invariant covariant derivative (21) can be expressed as*

$$\nabla_X \mathcal{Y} = \frac{1}{2} ([X, \mathcal{Y}] - \mathbb{I}^{-1} (\text{ad}_X^* \mathbb{I} \mathcal{Y} + \text{ad}_Y^* \mathbb{I} X)). \tag{27}$$

Proof. Using (24) to extract \mathcal{Z} on both terms on the right hand side expression (26) the result follows. □

A consequence is that the exponential map of the Lie group does not agree with the Riemannian exponential map corresponding to \mathbb{I} :

Proposition 4.3. *The one-parameter subgroups of G do not coincide with the geodesics of \mathbb{I} .*

Proof. The one-parameter subgroups of a Lie group like G correspond to the autotransported curves through the identity of an affine connection if and only if $\nabla_X X = 0$ for all the Lie algebra valued vectors X [14], Prop. 2.9, Chap. X). □

In fact, using the language of [14], $\nabla_X \mathcal{X} = 0$ means that the natural torsion-free and canonical connections can be made to have the same geodesics. The coincidence holds only when the bilinear map $U(\mathcal{X}, \mathcal{Y})$ above is compatible with the metric in the following way:

Proposition 4.4. *Given a positive definite Riemannian metric on a group manifold G :*

$$\nabla_X \mathcal{X} = 0 \quad \forall \mathcal{X} \in \mathfrak{g} \iff \langle U(\mathcal{X}, \mathcal{Y}), \mathcal{Z} \rangle = 0 \quad \forall \mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}.$$

Proof. The proof can be obtained from part 2 of Theorem 3.3 Chapter X in [14]. To see it directly use (25): assuming the right hand side expression holds true, if $\nabla_X \mathcal{X} \neq 0$ it has to be $U(\mathcal{X}, \mathcal{X}) \neq 0$ but then it is enough to choose a suitable \mathcal{Z} to obtain $\langle U(\mathcal{X}, \mathcal{Y}), \mathcal{Z} \rangle \neq 0$ for some $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ therefore getting into contradiction. Similarly, on the other direction, assume $\nabla_X \mathcal{X} = 0$. If $\langle U(\mathcal{X}, \mathcal{Y}), \mathcal{Z} \rangle \neq 0$ for some $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ then $U(\mathcal{X}, \mathcal{Y}) \neq 0$ and therefore also $U(\tilde{\mathcal{X}}, \tilde{\mathcal{X}}) \neq 0$ for some $\tilde{\mathcal{X}}$ because U is a symmetric bilinear map, which is again a contradiction. \square

The condition $\langle U(\mathcal{X}, \mathcal{Y}), \mathcal{Z} \rangle = 0$ for all $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}$ holds for the so-called *naturally reductive homogeneous spaces*. This is the case for example of compact Lie groups with bi-invariant metric: its covariant derivative is well-known to be simply $\nabla_X \mathcal{Y} = \frac{1}{2}[\mathcal{X}, \mathcal{Y}]$.

It is perhaps worth to make a further comment on the relation between ad_X , left invariance and Killing vectors. Since $\nabla_X Y = \text{ad}_X Y - A_X Y$, rewriting (6) as

$$\underbrace{\mathcal{L}_X \langle Y, Z \rangle}_{(*)} = \langle \text{ad}_X Y, Z \rangle + \langle Y, \text{ad}_X Z \rangle + \underbrace{\langle A_X Y, Z \rangle + \langle Y, A_X Z \rangle}_{(**)} \tag{28}$$

we have:

- (1) if Y, Z are left invariant, then $(*) = 0$;
- (2) if X is a Killing vector, $(**) = 0$;
- (3) if Y, Z are left invariant and X is a Killing vector (as in (23)), then ad_X is skew-symmetric with respect to $\langle \cdot, \cdot \rangle$, *i.e.*, the following equivalent quantities

$$\langle \text{ad}_X Y, Z \rangle + \langle Y, \text{ad}_X Z \rangle = 2\langle U(Y, Z), X \rangle = -\langle \mathbb{I}^{-1}(\text{ad}_Y^* \mathbb{I}Z + \text{ad}_Z^* \mathbb{I}Y), X \rangle \tag{29}$$

are all vanishing.

A Lie group admits a bi-invariant metric if and only if ad_X is skew-symmetric with respect to $\langle \cdot, \cdot \rangle$ for all $X \in \mathfrak{g}$ [20, Lem. 7.2). In our case, even considering left invariant vector fields, this is not the case as can be deduced from the expression of $U(\mathcal{Y}, \mathcal{Z})$ calculated in Proposition 4.2.

Given $X \in \mathfrak{g}$, $\mathcal{X} = a^i A_i$ call $\dot{\mathcal{X}} = \frac{\partial a^i}{\partial t} A_i$. The local coordinate chart at $\gamma \in G$ is given by left translating the time-1 Lie group exponential map of A_1, \dots, A_n : $x^i = \gamma e^{A_i}$ (so that a basis of tangent vectors at γ is indeed $\frac{\partial}{\partial x^i} = B_i = \gamma A_i$). Such coordinates are not Riemannian normal coordinates since the Christoffel symbols are nonnull. The covariant derivative of $Y = \gamma \mathcal{Y}$, $\mathcal{Y} = b^i A_i$, in the direction of $X = \gamma \mathcal{X} = \gamma a^i A_i$ becomes:

$$\begin{aligned} \nabla_X Y &= (\nabla_X Y)^k B_k = \left(a^i \frac{\partial b^k}{\partial x^i} + a^i b^j \Gamma_{ij}^k \right) B_k \\ &= ((\mathcal{L}_X b^k) + a^i b^j \Gamma_{ij}^k) \gamma A_k = \gamma (\mathcal{L}_X b^k) A_k + \gamma a^i b^j \nabla_{A_i} A_j && \text{by (21)} \\ &= \gamma ((\mathcal{L}_X b^k) A_k + \nabla_X \mathcal{Y}). \end{aligned} \tag{30}$$

Left invariance allows one to express vector fields on G and vector fields on TG by means of their pull-back to \mathfrak{g} with respect to the same basis *i.e.*, $\frac{\partial}{\partial x^i} = \gamma A_i$ and $\frac{\partial}{\partial v^i} = \gamma A_i$, $i = 1, \dots, n$. The parallel transport of any $Y = \gamma b^i A_i$ along γ gives rise to a horizontal lift of the curve $\dot{\gamma} = \gamma \mathcal{X}$ to the tangent bundle curve $\gamma^h = (\gamma, Y)$ having as tangent vector

$$X^h = \frac{d\gamma^h}{dt} = \gamma \mathcal{X}^h = (\gamma a^k A_k; -\gamma \Gamma_{ij}^k a^i b^j A_k)$$

(the first component living on $T_\gamma G$, the second on $T_{(\gamma, \gamma \mathfrak{y})} TG$). For matrix groups, $\nabla_{\dot{\gamma}} Y = 0$ corresponds to a linear frame being parallel transported along γ , see (16). Calling $\dot{X} = \frac{\partial b^i}{\partial t} A_i$, in general, from (25), the covariant derivative is decomposed in the three parts:

$$\gamma \left(\dot{\mathfrak{y}} + \frac{1}{2}[\mathfrak{X}, \mathfrak{y}] + U(\mathfrak{X}, \mathfrak{y}) \right) = 0.$$

Therefore a coordinate independent expression for X^h is

$$X^h = \left(\gamma \mathfrak{X}; -\gamma \left(\frac{1}{2}[\mathfrak{X}, \mathfrak{y}] + U(\mathfrak{X}, \mathfrak{y}) \right) \right) \tag{31}$$

whose integral curves are

$$\begin{aligned} \dot{\gamma} &= \gamma \mathfrak{X} \\ \gamma \dot{\mathfrak{y}} &= -\gamma \nabla_x \mathfrak{y} = -\gamma \left(\frac{1}{2}[\mathfrak{X}, \mathfrak{y}] + U(\mathfrak{X}, \mathfrak{y}) \right). \end{aligned}$$

4.1. The fiber bundle picture for group symmetries

Consider the tangent bundle TG of the n -dimensional Lie group G . For each $g \in G$, the fiber $\tau^{-1}(g)$ of this tangent bundle is the tangent space at g , $T_g G$, isomorphic to \mathbb{R}^n . Left invariance gives to the tangent bundle the structure of principal fiber bundle with structure group G and base manifold \mathfrak{g} , by using the isomorphism between $T_g G$ and $T_e G = \mathfrak{g}$ that left (or right) translation implies. The three properties of a principal fiber bundle (see [14], Ch. II, p. 50) are trivial to verify for $TG(\mathfrak{g}, G)$. In fact, \mathfrak{g} is the quotient space of TG by the equivalence relation induced by G , the projection $\pi : TG \rightarrow \mathfrak{g}$ is the left translation itself and the fibers $\pi^{-1}(\mathfrak{X})$, $\mathfrak{X} \in \mathfrak{g}$, are isomorphic to G since the left invariant action is free and transitive. Furthermore, by considering left invariant vector fields, TG is made into a globally trivial fiber bundle *via* the map $T_g G \rightarrow G \times \mathfrak{g}$, $(g, v_g) \mapsto (g, L_{g^{-1}*} v_g)$. For a generic smooth manifold, there always exists a principal fiber bundle similar to the one considered here and it is the frame bundle $GL_n(\mathbb{R})$ obtained by all possible linear changes of basis of the tangent space (isomorphic to \mathbb{R}^n) at any point of the manifold. In particular, the fiber bundle we are considering is obtained by “reducing” $GL_n(\mathbb{R})$ to its subgroup G and is normally referred to as G -structure.

From (16), the condition $\nabla_{\dot{\gamma}} Y = 0$ allows one to describe vectors fields of $\mathcal{D}(G)$ that are horizontal in the tangent bundle with respect to ∇ and the lifting procedure described in Section 2.2. From Propositions 4.3 and 4.4, the compatibility condition between horizontal curves of \mathbb{I} and horizontal curves of the fiber bundle structure of G (*i.e.*, the “ G -structure” described above) is that ∇ invariant to left translations and that $\nabla_x \mathfrak{X} = 0$. In fact, invariance plus $\nabla_x \mathfrak{X} = 0$ means that the parallel displacement can be carried out by left translations regardless of the path to follow.

The appearing of a nonvanishing term $U(\mathfrak{X}, \mathfrak{y})$ implies that the reduction process produces a “geometric phase”, *i.e.*, out of horizontal vector fields (in the tangent bundle) one also obtains a vertical vector field (in the fiber bundle). From (25), this happens whenever $\nabla_x \mathfrak{y}$ is not completely skew symmetric.

So for the horizontal lift (31) the $T_{(\gamma, \gamma \mathfrak{y})} TG$ term splits into the horizontal component (in the fiber bundle) $-\gamma \frac{1}{2}[\mathfrak{X}, \mathfrak{y}]$ and the vertical one $-\gamma U(\mathfrak{X}, \mathfrak{y})$.

5. REDUCTION OF HAMILTON PRINCIPLE BY GROUP INVARIANCE

The variational principle as stated in Section 2.1 holds for generic Riemannian manifolds and does not take advantage of the group structure of G . In particular, the left invariance properties of a Lie group allow to reduce the infinitesimal variations from the tangent bundle to the corresponding Lie algebra: this subject is treated extensively in the book [19]. Also the semidirect product structure of G can be exploited explicitly in what is

called reduction by stages, especially when V has nonzero fixed points under K and the corresponding isotropy subgroup is of particular interest, like in the heavy top case [8].

Left invariance allows one to express the vector field $\dot{\gamma} \in TG$ in terms of its pullback to the identity: $L_{\gamma}^* \dot{\gamma} = \gamma^{-1} \dot{\gamma} \in \mathfrak{g}$. Consider proper variations $\mathcal{G}(s, t)$ of $\gamma(t)$ with tangent bundle infinitesimal variations $S(s, t) = \frac{d}{ds} \mathcal{G}^{(t)}(s)$ and $T(s, t) = \frac{d}{dt} \mathcal{G}_s(t)$, along the main and transverse curves respectively. Call $\mathcal{T}(t)$ and $\mathcal{S}(t)$ the \mathfrak{g} -valued infinitesimal variations corresponding to $\dot{\gamma}$ and $\delta\gamma$. Since $\mathcal{G}_0(t) = \gamma(t)$, they are uniquely defined by the two relations:

$$\begin{aligned} T(0, t) &= \dot{\gamma}(t) = \mathcal{G}_0(t)\mathcal{T}(t) = \gamma(t)\mathcal{T}(t) \\ S(0, t) &= \delta\gamma(t) = \mathcal{G}_0(t)\mathcal{S}(t) = \gamma(t)\mathcal{S}(t). \end{aligned} \tag{32}$$

We need to compute the covariant derivatives $\nabla_{\delta\gamma}\dot{\gamma}$ and $\nabla_{\dot{\gamma}}\delta\gamma$ in terms of \mathcal{T} and \mathcal{S} . Considering the basis A_1, \dots, A_n of \mathfrak{g} , in coordinates $\mathcal{T}(t) = \alpha^i(t)A_i$ and $\mathcal{S}(t) = \beta^i(t)A_i$. Call $\dot{\mathcal{T}} = \frac{\partial\alpha^i}{\partial t}A_i$ and $\mathcal{T}' = \frac{\partial\alpha^i}{\partial s}A_i$ and similarly for \mathcal{S} . The coordinate functions α^i and β^i are defined along the family of curves \mathcal{G} . In particular, from (11), along the main and transverse curves $\mathcal{G}_s(t)$ and $\mathcal{G}^{(t)}(s)$ the Lie derivatives $\mathcal{L}_T \cdot$ and $\mathcal{L}_S \cdot$ becomes derivatives in t and s respectively. Therefore $\mathcal{L}_{\dot{\gamma}}\alpha^k = \frac{\partial\alpha^k}{\partial t}$ and $\mathcal{L}_{\delta\gamma}\alpha^k = \frac{\partial\alpha^k}{\partial s}$ and similarly for β^k . Therefore (30) becomes the following:

Proposition 5.1. *Consider the Lie group G with left-invariant Riemannian connection ∇ . For the proper variations $\mathcal{G}(s, t)$, the covariant derivatives $\nabla_{\dot{\gamma}}\delta\gamma$ and $\nabla_{\delta\gamma}\dot{\gamma}$ have the following left-invariant expressions:*

$$\begin{aligned} \nabla_{\dot{\gamma}}\delta\gamma &= \gamma \left(\dot{\mathcal{S}} + \nabla_{\mathcal{T}}\mathcal{S} \right) \\ \nabla_{\delta\gamma}\dot{\gamma} &= \gamma \left(\mathcal{T}' + \nabla_{\mathcal{S}}\mathcal{T} \right). \end{aligned} \tag{33}$$

Proof.

$$\begin{aligned} \nabla_{\dot{\gamma}}\delta\gamma &= \nabla_{\gamma\mathcal{T}(t)}\gamma\mathcal{S}(t) = \nabla_{\alpha^i(t)B_i}\beta^j(t)B_j \\ &= (\mathcal{L}_{\dot{\gamma}}\beta^j) B_j + \alpha^i\beta^j\nabla_{B_i}B_j = \frac{\partial\beta^j}{\partial t}B_j + \gamma\alpha^i\beta^j\nabla_{A_i}A_j && \text{by (21)} \\ &= \gamma\frac{\partial\beta^j}{\partial t}A_j + \frac{1}{2}\gamma\alpha^i\beta^j \left([A_i, A_j] - \mathbb{I}^{-1}\text{ad}_{A_i}^*\mathbb{I}A_j - \mathbb{I}^{-1}\text{ad}_{A_j}^*\mathbb{I}A_i \right) \\ &= \gamma\frac{\partial\beta^j}{\partial t}A_j + \frac{1}{2}\gamma \left([\mathcal{T}, \mathcal{S}] - \mathbb{I}^{-1}\text{ad}_{\mathcal{T}}^*\mathbb{I}\mathcal{S} - \mathbb{I}^{-1}\text{ad}_{\mathcal{S}}^*\mathbb{I}\mathcal{T} \right) = \gamma \left(\dot{\mathcal{S}} + \nabla_{\mathcal{T}}\mathcal{S} \right) \end{aligned}$$

and similarly

$$\nabla_{\delta\gamma}\dot{\gamma} = \nabla_{\gamma\mathcal{S}(t)}\gamma\mathcal{T}(t) = \gamma\frac{\partial\alpha^i}{\partial s}A_i + \frac{1}{2}\gamma \left([\mathcal{S}, \mathcal{T}] - \mathbb{I}^{-1}\text{ad}_{\mathcal{S}}^*\mathbb{I}\mathcal{T} - \mathbb{I}^{-1}\text{ad}_{\mathcal{T}}^*\mathbb{I}\mathcal{S} \right) = \gamma \left(\mathcal{T}' + \nabla_{\mathcal{S}}\mathcal{T} \right). \quad \square$$

The presence of two terms in both the covariant derivatives (33) is due to the affine nature of the connection. The terms \mathcal{T}' or $\dot{\mathcal{S}}$ appear when the covariant derivative is calculated out of the identity of the group. Here and in the following the covariant derivatives involving \mathcal{T} or \mathcal{S} are always calculated in the identity element of the group.

Lemma 5.2 (symmetry lemma for reduction by group invariance in the Riemannian case). *In the case of group manifold with Levi-Civita connection, the Lie algebra valued mixed derivatives are related by*

$$\mathcal{T}' = \dot{\mathcal{S}} \tag{34}$$

where \mathcal{T} and \mathcal{S} are respectively the Lie algebra valued infinitesimal variations for $\dot{\gamma}$ and $\delta\gamma$ defined in (32). Furthermore

$$\nabla_{\mathcal{T}}\mathcal{S} = \nabla_{\mathcal{S}}\mathcal{T}. \tag{35}$$

Proof. From $\nabla_{\dot{\gamma}}\delta\gamma = \nabla_{\delta\gamma}\dot{\gamma}$ and the expression (33) for the two covariant derivatives:

$$\mathcal{T}' = \dot{\mathcal{S}} + [\mathcal{T}, \mathcal{S}]. \tag{36}$$

But using left invariance: $[\mathcal{T}, \mathcal{S}] = \gamma^{-1}[\gamma\mathcal{T}, \gamma\mathcal{S}] = \gamma^{-1}[\dot{\gamma}, \delta\gamma] = 0$ by the symmetry lemma. Consequently also (35) follows. \square

Notice that this is true only because we have chosen a torsion-free connection. In general when a different connection is chosen on the Lie group (for example the (+) or (-) canonical connections of Cartan [7]), the “covariant infinitesimal variations” for the reduced principle have the more general expression (36), see [18] and references therein.

In general, it is not possible to conclude on \mathcal{T} and \mathcal{S} being Killing vector fields without knowing the metric tensor \mathbb{I} . By definition of Levi-Civita connection, the parallel transport along γ leaves the inner product invariant. An equation like (22), corresponding to parallel transport of the inner product along γ for left invariant vector fields, splits after the reduction into two types of Lie algebra valued infinitesimal covariant variations, those indicated by “ $\dot{\cdot}$ ” (or “ $\dot{\prime}$ ”) and those by the covariant derivative symbol. For $\mathcal{Y}, \mathcal{Z} \in \mathfrak{g}$ along the curve γ of tangent vector $\dot{\gamma} = \gamma\mathcal{T}$, abusing notation one could write:

$$\frac{d}{dt}\langle \mathcal{Y}, \mathcal{Z} \rangle = \langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle + \nabla_{\mathcal{T}}\langle \mathcal{Y}, \mathcal{Z} \rangle = 0$$

where $\langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle = \langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \dot{\mathcal{Z}} \rangle$ and $\nabla_{\mathcal{T}}\langle \mathcal{Y}, \mathcal{Z} \rangle = \langle \nabla_{\mathcal{T}}\mathcal{Y}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \nabla_{\mathcal{T}}\mathcal{Z} \rangle$ are, respectively, the affine part and the linear part of the parallel transported inner product along γ . This complicates the expression for the reduced equations as none of the infinitesimal variations alone is Killing. However, the following proposition shows that each of them respects (22).

Proposition 5.3. *Given $\mathcal{T}, \mathcal{Y}, \mathcal{Z} \in \mathfrak{g}$, the equation (22) for parallel transport of the inner product along a curve $\gamma \in G$ with tangent vector $T = \gamma\mathcal{T}$ splits into the two relations:*

$$\langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle = -\langle \mathcal{Y}, \dot{\mathcal{Z}} \rangle \tag{37}$$

$$\langle \nabla_{\mathcal{T}}\mathcal{Y}, \mathcal{Z} \rangle = -\langle \mathcal{Y}, \nabla_{\mathcal{T}}\mathcal{Z} \rangle. \tag{38}$$

Proof. Straightforward calculation from (12):

$$\begin{aligned} \frac{d}{dt}\langle \gamma\mathcal{Y}, \gamma\mathcal{Z} \rangle &= \langle \nabla_{\gamma\mathcal{T}}\gamma\mathcal{Y}, \gamma\mathcal{Z} \rangle + \langle \gamma\mathcal{Y}, \nabla_{\gamma\mathcal{T}}\gamma\mathcal{Z} \rangle \\ &= \langle \gamma(\dot{\mathcal{Y}} + \nabla_{\mathcal{T}}\mathcal{Y}), \gamma\mathcal{Z} \rangle + \langle \gamma\mathcal{Y}, \gamma(\dot{\mathcal{Z}} + \nabla_{\mathcal{T}}\mathcal{Z}) \rangle \end{aligned}$$

i.e.,

$$\begin{aligned} \frac{d}{dt}\langle \mathcal{Y}, \mathcal{Z} \rangle &= \langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \dot{\mathcal{Z}} \rangle + \langle \nabla_{\mathcal{T}}\mathcal{Y}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \nabla_{\mathcal{T}}\mathcal{Z} \rangle \\ 0 &= \langle \dot{\mathcal{Y}}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \dot{\mathcal{Z}} \rangle \end{aligned}$$

from left invariance of the metric and (22). \square

Corollary 5.4. *For the family of variations $\mathcal{G}(s, t)$ of γ , with the notation above,*

$$\langle \dot{\mathcal{S}}, \mathcal{T} \rangle = -\langle \mathcal{S}, \dot{\mathcal{T}} \rangle \tag{39}$$

$$\langle \nabla_{\mathcal{T}} \mathcal{S}, \mathcal{T} \rangle = -\langle \nabla_{\mathcal{T}} \mathcal{T}, \mathcal{S} \rangle \tag{40}$$

$$\langle \nabla_{\mathcal{S}} \mathcal{S}, \mathcal{T} \rangle = -\langle \mathcal{S}, \nabla_{\mathcal{S}} \mathcal{T} \rangle. \tag{41}$$

When the homogeneous space is naturally reductive, $\nabla_{\mathcal{T}} \mathcal{T} = 0$ implies in (40) that also $\langle \nabla_{\mathcal{T}} \mathcal{S}, \mathcal{T} \rangle = 0$ and therefore the situation is much simpler.

Remark 5.5. $\nabla_{\mathcal{T}} \mathcal{S} = \nabla_{\mathcal{S}} \mathcal{T}$ is symmetric and belongs to V . In fact, from (27) and $[\mathcal{T}, \mathcal{S}] = 0$

$$\nabla_{\mathcal{T}} \mathcal{S} = \nabla_{\mathcal{S}} \mathcal{T} = U(\mathcal{S}, \mathcal{T}) = -\frac{1}{2} \mathbb{I}^{-1} (\text{ad}_{\mathcal{S}}^* \mathbb{I} \mathcal{T} + \text{ad}_{\mathcal{T}}^* \mathbb{I} \mathcal{S}).$$

Another consequence of $[\mathcal{T}, \mathcal{S}] = 0$ is the following:

Remark 5.6. For the reduced infinitesimal variations of $\mathcal{G}(s, t)$, the covariant derivatives are vertical in the fiber bundle.

From the reduced symmetry lemma we obtain the Euler-Poincaré equations. The result is well-known (see [19]) although it is normally not obtained using exclusively the tools from Riemannian geometry as we do here.

Theorem 5.7 (reduced Hamilton principle). *For (G, \mathbb{I}) , the critical curves of the left invariant energy functional $\mathcal{E} = \frac{1}{2} \int_{t_0}^{t_f} \langle \mathcal{T}, \mathcal{T} \rangle dt$, where $\mathcal{T} = \gamma^{-1}(t) \dot{\gamma}(t)$, in correspondence of proper variations $\mathcal{G}(s, t)$ (and of their “covariant infinitesimal variations”), are given by the Euler-Poincaré equations*

$$\dot{\mathcal{T}} = -\nabla_{\mathcal{T}} \mathcal{T} = \mathbb{I}^{-1} \text{ad}_{\mathcal{T}}^* \mathbb{I} \mathcal{T}. \tag{42}$$

Proof. The proof can be obtained directly by inserting into the first variation formula (13) the value of the covariant derivative (33). Likewise, going through the reduction of the functional \mathcal{E} :

$$\begin{aligned} \left. \frac{d}{ds} \mathcal{E}(\mathcal{G}_s(t)) \right|_{s=0} &= \left. \int_{t_0}^{t_f} \langle \nabla_S T, T \rangle dt \right|_{s=0} = \int_{t_0}^{t_f} \langle \mathcal{T}' + \nabla_{\mathcal{S}} \mathcal{T}, \mathcal{T} \rangle dt \\ &= - \int_{t_0}^{t_f} \langle \dot{\mathcal{T}}, \mathcal{S} \rangle dt - \int_{t_0}^{t_f} \langle \nabla_{\mathcal{T}} \mathcal{T}, \mathcal{S} \rangle dt && \text{by (39) and (40)} \\ &= - \int_{t_0}^{t_f} \langle \dot{\mathcal{T}} + \nabla_{\mathcal{T}} \mathcal{T}, \mathcal{S} \rangle dt. \end{aligned}$$

□

The geodesic spray whose integral curves are the Euler-Poincaré equations (42) is

$$\Gamma = \gamma \mathcal{G} = \gamma(\mathcal{T}; -U(\mathcal{T}, \mathcal{T})).$$

The component in $T_{(\gamma, \gamma \mathcal{T})} TG$ is purely symmetric and therefore it is vertical in the fiber bundle. In fact, it disappears on naturally reductive homogeneous spaces, where the reduced Euler-Lagrange equations (42) have only a left hand side: $\dot{\gamma} = \gamma \mathcal{T}$, $\dot{\mathcal{T}} = 0$.

6. REDUCTION OF THE SECOND ORDER VARIATIONAL PROBLEM

For $G = K \otimes V$, we want to compute the reduction by group symmetry of the necessary condition (19) of Theorem 3.1.

When reducing higher order covariant derivatives, the two components of the covariant derivative mentioned in Section 5 mix up. For example, the “ \cdot ” part, applied to $\nabla_x \mathcal{Y}$ results into:

$$(\nabla_x \dot{\mathcal{Y}}) = \frac{\partial a^i b^j}{\partial t} \nabla_{A_i} A_j = \left(\frac{\partial a^i}{\partial t} b^j + a^i \frac{\partial b^j}{\partial t} \right) \nabla_{A_i} A_j = \nabla_{\dot{x}} \mathcal{Y} + \nabla_x \dot{\mathcal{Y}}. \quad (43)$$

A few useful relations are:

Proposition 6.1. *For the family of variations $\mathcal{G}(s, t)$ on (G, \mathbb{I}) :*

$$\nabla_{\delta\gamma} \nabla_{\dot{\gamma}} \dot{\gamma} = \gamma \left((\dot{\mathcal{J}})' + \nabla_{\mathcal{S}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{S}}} \mathcal{J} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J} \right) \quad (44)$$

$$\nabla_{\dot{\gamma}} \nabla_{\delta\gamma} \dot{\gamma} = \gamma \left((\dot{\mathcal{J}})' + \nabla_{\mathcal{S}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{S}}} \mathcal{J} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_{\mathcal{T}} \nabla_{\mathcal{S}} \mathcal{J} \right) \quad (45)$$

$$\nabla^2_{\dot{\gamma}} \dot{\gamma} = \gamma \left(\ddot{\mathcal{J}} + 2\nabla_{\mathcal{T}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{T}}} \mathcal{J} + \nabla^2_{\mathcal{T}} \mathcal{J} \right) \quad (46)$$

$$\nabla^3_{\dot{\gamma}} \dot{\gamma} = \gamma \left(\ddot{\mathcal{J}} + 3\nabla_{\mathcal{T}} \ddot{\mathcal{J}} + 3\nabla_{\dot{\mathcal{T}}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{T}}} \mathcal{J} + 3\nabla^2_{\mathcal{T}} \dot{\mathcal{J}} + \right. \quad (47)$$

$$\left. + 2\nabla_{\mathcal{T}} \nabla_{\dot{\mathcal{T}}} \mathcal{J} + \nabla_{\dot{\mathcal{T}}} \nabla_{\mathcal{T}} \mathcal{J} + \nabla^3_{\mathcal{T}} \mathcal{J} \right)$$

$$R(\mathcal{S}, \mathcal{T}) \mathcal{T} = \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{T} - \nabla_{\mathcal{T}} \nabla_{\mathcal{S}} \mathcal{T} = [\nabla_{\mathcal{S}}, \nabla_{\mathcal{T}}] \mathcal{T}. \quad (48)$$

Proof. We only prove (44), the other calculations being similar.

$$\begin{aligned} \nabla_{\delta\gamma} \nabla_{\dot{\gamma}} \dot{\gamma} &= \nabla_{\delta\gamma} \left(\gamma \left(\dot{\mathcal{J}} + \nabla_{\mathcal{T}} \mathcal{J} \right) \right) \\ &= \mathcal{L}_{\delta\gamma} \left(\frac{\partial \alpha^j}{\partial t} \right) \gamma A_j + \beta^i \frac{\partial \alpha^j}{\partial t} \gamma \nabla_{A_i} A_j + \mathcal{L}_{\delta\gamma} (\alpha^i \alpha^j) \nabla_{A_i} A_j + \gamma \alpha^i \alpha^j \beta^k \nabla_{A_k} \nabla_{A_i} A_j \\ &= \gamma \left(\frac{\partial^2 \alpha^j}{\partial s \partial t} A_j + \left(\beta^i \frac{\partial \alpha^j}{\partial t} + \frac{\partial \alpha^i}{\partial t} \alpha^j + \alpha^i \frac{\partial \alpha^j}{\partial t} \right) \nabla_{A_i} A_j + \beta^k \nabla_{A_k} \nabla_{\mathcal{T}} \mathcal{J} \right) \\ &= \gamma \left((\dot{\mathcal{J}})' + \nabla_{\mathcal{S}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{S}}} \mathcal{J} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J} \right). \end{aligned}$$

Concerning (48), from (10) with $[\delta\gamma, \dot{\gamma}] = [\mathcal{S}, \mathcal{T}] = 0$, from (44) and (45)

$$\begin{aligned} R(\delta\gamma, \dot{\gamma}) \dot{\gamma} &= \nabla_{\delta\gamma} \nabla_{\dot{\gamma}} \dot{\gamma} - \nabla_{\dot{\gamma}} \nabla_{\delta\gamma} \dot{\gamma} \\ R(\gamma \mathcal{S}, \gamma \mathcal{T}) \gamma \mathcal{T} &= \gamma \left((\dot{\mathcal{J}})' + \nabla_{\mathcal{S}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{S}}} \mathcal{J} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J} - \right. \\ &\quad \left. - (\dot{\mathcal{J}})' - \nabla_{\mathcal{S}} \dot{\mathcal{J}} - \nabla_{\dot{\mathcal{S}}} \mathcal{J} - \nabla_{\mathcal{T}} \dot{\mathcal{S}} - \nabla_{\mathcal{T}} \nabla_{\mathcal{S}} \mathcal{J} \right). \end{aligned}$$

Since R is a tensor it is left invariant; furthermore, the order of the mixed second derivative with respect to s and t commutes also in \mathfrak{g}

$$(\dot{\mathcal{J}})' = \frac{\partial^2 \alpha^j}{\partial s \partial t} A_j = \frac{\partial^2 \alpha^j}{\partial t \partial s} A_j = (\dot{\mathcal{J}})'. \quad (49)$$

Hence, the result. \square

The left invariance of the curvature tensor R means that the curvature term of (19) is:

$$R(\nabla_{\dot{\gamma}} \dot{\gamma}, \dot{\gamma}) \dot{\gamma} = R(\dot{\mathcal{J}} + \nabla_{\mathcal{T}} \mathcal{J}, \mathcal{T}) \gamma \mathcal{T} = \gamma \left(R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{T} + R(\nabla_{\mathcal{T}} \mathcal{J}, \mathcal{T}) \mathcal{T} \right). \quad (50)$$

The expressions (47) and (50) allows one to write down directly the reduced expression for (19) in terms of left invariant vector fields. However, it is quite instructive to see the genesis of this formula, going through the reduction of the cost functional \mathcal{J} . We compute first the following equalities:

Proposition 6.2. *For the family of variations $\mathcal{G}(s, t)$ on (G, \mathbb{I}) :*

$$\langle (\dot{\mathcal{J}})', \dot{\mathcal{J}} \rangle = \langle \mathcal{S}, \ddot{\mathcal{J}} \rangle \quad (51)$$

$$\langle (\dot{\mathcal{J}})', \nabla_{\mathcal{T}} \mathcal{J} \rangle = \langle \mathcal{S}, \nabla_{\dot{\mathcal{J}}} \mathcal{J} + 2\nabla_{\dot{\mathcal{J}}} \dot{\mathcal{J}} + \nabla_{\mathcal{T}} \ddot{\mathcal{J}} \rangle \quad (52)$$

$$\langle \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J}, \dot{\mathcal{J}} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}}^2 \dot{\mathcal{J}} + R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{J} \rangle \quad (53)$$

$$\langle \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J}, \nabla_{\mathcal{T}} \mathcal{J} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}}^3 \mathcal{J} + R(\nabla_{\mathcal{T}} \mathcal{J}, \mathcal{T}) \mathcal{J} \rangle \quad (54)$$

$$\langle \nabla_{\mathcal{T}} \dot{\mathcal{S}}, \dot{\mathcal{J}} \rangle = \langle \mathcal{S}, \nabla_{\dot{\mathcal{J}}} \dot{\mathcal{J}} + \nabla_{\mathcal{T}} \ddot{\mathcal{J}} \rangle \quad (55)$$

$$\langle \nabla_{\mathcal{T}} \dot{\mathcal{S}}, \nabla_{\mathcal{T}} \mathcal{J} \rangle = \langle \mathcal{S}, \nabla_{\dot{\mathcal{J}}} \nabla_{\mathcal{T}} \mathcal{J} + \nabla_{\mathcal{T}} \nabla_{\dot{\mathcal{J}}} \mathcal{J} + \nabla_{\mathcal{T}}^2 \dot{\mathcal{J}} \rangle \quad (56)$$

$$\langle \nabla_{\mathcal{S}} \dot{\mathcal{J}}, \dot{\mathcal{J}} \rangle = 0 \quad (57)$$

$$\langle \nabla_{\mathcal{S}} \dot{\mathcal{J}}, \nabla_{\mathcal{T}} \mathcal{J} \rangle = -\langle \dot{\mathcal{J}}, \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J} \rangle \quad (58)$$

$$\langle \nabla_{\dot{\mathcal{S}}} \mathcal{J}, \dot{\mathcal{J}} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}} \ddot{\mathcal{J}} \rangle \quad (59)$$

$$\langle \nabla_{\dot{\mathcal{S}}} \mathcal{J}, \nabla_{\mathcal{T}} \mathcal{J} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}} \nabla_{\dot{\mathcal{J}}} \mathcal{J} + 2\nabla_{\mathcal{T}}^2 \dot{\mathcal{J}} + R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{J} \rangle. \quad (60)$$

Proof. All the expressions are based on the equations (37) and (38). We see some of the significant calculations:

- Equation (51):

$$\langle (\dot{\mathcal{J}})', \dot{\mathcal{J}} \rangle = \langle (\dot{\mathcal{J}})', \dot{\mathcal{J}} \rangle = -\langle \mathcal{T}', \ddot{\mathcal{J}} \rangle = -\langle \dot{\mathcal{S}}, \ddot{\mathcal{J}} \rangle = \langle \mathcal{S}, \ddot{\mathcal{J}} \rangle \quad \text{by (49);}$$

- Equation (52):

$$\begin{aligned} \langle (\dot{\mathcal{J}})', \nabla_{\mathcal{T}} \mathcal{J} \rangle &= \langle \mathcal{S}, (\nabla_{\mathcal{T}} \ddot{\mathcal{J}}) \rangle = \langle \mathcal{S}, \frac{\partial^2 \alpha^i \alpha^j}{\partial t^2} \nabla_{A_i} A_j \rangle \\ &= \langle \mathcal{S}, \left(\frac{\partial^2 \alpha^i}{\partial t^2} \alpha^j + 2 \frac{\partial \alpha^i}{\partial t} \frac{\partial \alpha^j}{\partial t} + \alpha^i \frac{\partial^2 \alpha^j}{\partial t^2} \right) \nabla_{A_i} A_j \rangle \\ &= \langle \mathcal{S}, \nabla_{\dot{\mathcal{J}}} \mathcal{J} + 2\nabla_{\dot{\mathcal{J}}} \dot{\mathcal{J}} + \nabla_{\mathcal{T}} \ddot{\mathcal{J}} \rangle; \end{aligned}$$

- Equation (53):

$$\begin{aligned} \langle \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{J}, \dot{\mathcal{J}} \rangle &= \langle \nabla_{\mathcal{T}} \nabla_{\mathcal{S}} \mathcal{J} + R(\mathcal{S}, \mathcal{T}) \mathcal{J}, \dot{\mathcal{J}} \rangle \\ &= -\langle \nabla_{\mathcal{S}} \mathcal{J}, \nabla_{\mathcal{T}} \dot{\mathcal{J}} \rangle + \langle R(\mathcal{S}, \mathcal{T}) \mathcal{J}, \dot{\mathcal{J}} \rangle \quad \text{by (48)} \\ &= -\langle \nabla_{\mathcal{T}} \mathcal{S}, \nabla_{\mathcal{T}} \dot{\mathcal{J}} \rangle + \langle R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{J}, \mathcal{S} \rangle \quad \text{by (35) and (20)} \\ &= \langle \mathcal{S}, \nabla_{\mathcal{T}}^2 \dot{\mathcal{J}} \rangle + \langle R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{J}, \mathcal{S} \rangle; \end{aligned}$$

- Equation (59):

$$\begin{aligned} \langle \nabla_{\dot{\mathcal{S}}} \mathcal{J}, \dot{\mathcal{J}} \rangle &= \langle (\nabla_{\mathcal{S}} \dot{\mathcal{J}}) - \nabla_{\mathcal{S}} \dot{\mathcal{J}}, \dot{\mathcal{J}} \rangle = -\langle \nabla_{\mathcal{S}} \mathcal{J}, \ddot{\mathcal{J}} \rangle - \langle \nabla_{\mathcal{S}} \dot{\mathcal{J}}, \dot{\mathcal{J}} \rangle \quad \text{by (43)} \\ &= -\langle \nabla_{\mathcal{T}} \mathcal{S}, \ddot{\mathcal{J}} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}} \ddot{\mathcal{J}} \rangle \quad \text{by (35) and (57)}. \end{aligned}$$

The other relations are obtained using similar arguments. \square

Theorem 6.3. *A necessary condition for a smooth curve $\gamma(t) \in G$ of tangent vector field $\mathcal{J} \in \mathfrak{g}$ and interpolating $\gamma(t_0) = g_0$, $\gamma(t_f) = g_f$, $V_0 = g_0^{-1} v_0$, $V_f = g_f^{-1} v_f$ to be an extremum of \mathcal{J} is that*

$$\ddot{\mathcal{J}} + 3\nabla_{\mathcal{T}} \dot{\mathcal{J}} + 3\nabla_{\dot{\mathcal{J}}} \dot{\mathcal{J}} + \nabla_{\dot{\mathcal{J}}} \mathcal{J} + 3\nabla_{\mathcal{T}}^2 \dot{\mathcal{J}} + 2\nabla_{\mathcal{T}} \nabla_{\dot{\mathcal{J}}} \mathcal{J} + \nabla_{\dot{\mathcal{J}}} \nabla_{\mathcal{T}} \mathcal{J} + \nabla_{\mathcal{T}}^3 \mathcal{J} + R(\dot{\mathcal{J}}, \mathcal{T}) \mathcal{J} + R(\nabla_{\mathcal{T}} \mathcal{J}, \mathcal{T}) \mathcal{J} = 0. \quad (61)$$

Proof. From (47) and (50) we have already (61) by brute force. To see it directly, expand the expression for \mathcal{J} and use the computations of Proposition 6.2.

$$\begin{aligned} \frac{d}{ds} \mathcal{J}(\mathcal{G}_s(t)) \Big|_{s=0} &= \int_{t_0}^{t_f} \langle \nabla_{\delta\gamma} \nabla_{\dot{\gamma}} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle dt \\ &= \int_{t_0}^{t_f} \langle \gamma \left((\dot{\mathcal{J}})' + \nabla_s \dot{\mathcal{J}} + \nabla_{\dot{s}} \mathcal{J} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_s \nabla_{\mathcal{T}} \mathcal{J} \right), \gamma \left(\dot{\mathcal{J}} + \nabla_{\mathcal{T}} \mathcal{J} \right) \rangle dt \end{aligned} \tag{44}$$

The ten inner products under the sign of integral are given in (51)–(60). The result follows by considering extremals of \mathcal{J} i.e., $\frac{d}{ds} \mathcal{J}(\mathcal{G}_s(t)) \Big|_{s=0} = 0$. \square

Notice that, except for pulling back the velocities v_0 and v_f to \mathfrak{g} , the boundary data on the curve itself g_0, g_f are not anymore entering into the problem, just like in the Euler-Poincaré equations.

7. APPLICATION TO OPTIMAL CONTROL OF MECHANICAL SYSTEMS

Recall that interpreting \mathbb{I} as the quadratic form of a kinetic energy, equation (3) corresponds to the Euler-Lagrange equation of a mechanical system having Lagrangian equal to kinetic energy. In this section we consider a forced Euler-Lagrange equation and use the material of the previous sections to study the reduction of an optimal control problem known in the literature as the “ C^2 dynamical interpolation problem”, see [10].

7.1. Simple mechanical control systems

If we add a forcing term to (3), we obtain a so-called *simple mechanical control system* [16] (without potential):

$$\nabla_{\dot{\gamma}} \dot{\gamma} = F(\gamma) \tag{62}$$

where $F = (F_1, \dots, F_n)$ is the control input distribution of \mathcal{M} . The vector fields $F_i = F_i(\gamma)$ are obtained by lowering the indices of the covectors \tilde{F}_i physically representing the forces or torques applied to the system: $F_i = \mathbb{I}^{-1} \tilde{F}_i$. Assuming F_1, \dots, F_n to be linearly independent on \mathcal{M} , then we have a *fully actuated mechanical system*. The system of first order differential equations corresponding to (62) was shown in [17] to be given by the second order vector field on $T\mathcal{M}$ obtained from the geodesic spray plus the vertical lifts of the input distribution:

$$\Gamma + F^v \tag{63}$$

having integral curves

$$\begin{aligned} \dot{x}^k &= v^k \\ \dot{v}^k &= -\Gamma_{ij}^k v^i v^j + F^k. \end{aligned}$$

From a control theory point of view, Γ is the drift of the system of first order differential equations and $F^v = (0 \frac{\partial}{\partial x^k} + F^k \frac{\partial}{\partial v^k})$ is the corresponding input vector field.

7.2. A C^2 dynamical interpolation problem

The presence of control inputs allows one to force the mechanical system along any suitable (feasible) trajectory, not necessarily those satisfying Hamilton principle of least action but rather a user or task defined cost functional. The condition (19) replaces the geodesic condition (3) in the sense that it superimposes a minimum acceleration motion to the “natural” geodesic motion associated with \mathbb{I} .

The resulting trajectory is C^∞ on \mathcal{M} and furthermore, by matching initial conditions of a new interval with the terminal data of the previous one, C^1 piecewise smooth trajectories on \mathcal{M} can be obtained. These are particularly useful for second order control systems as they represent the simplest curves feasible under

the ordinary assumption of piecewise continuous, measurable control inputs, generalization to a Riemannian manifold of Euclidean cubic splines. Equation (62) provides the expression for the control corresponding to a solution of (19). The full actuation of the mechanical control systems is a sufficient condition for free feasibility of the trajectories of (19).

If we restrict the class of control inputs to continuous piecewise smooth functions, then we have the above mentioned C^2 dynamical interpolation problem defined in [10]. A necessary condition for it can be obtained from Theorem 3.1 by choosing different boundary conditions: g_0, g_f, v_0 and $F(t_0) = \nabla_{\dot{\gamma}}\dot{\gamma}|_{\gamma=g_0, \dot{\gamma}=v_0}$. This last condition replaces the terminal boundary data v_f of Theorem 3.1. Using (62) to insert the control input F into (19), one obtains an equation that looks exactly like the Jacobi equation for F . However, the curve γ in this case is not a geodesic, instead it has to be computed together with the control action. Hence, what is used in the proposition below is *not* the Jacobi equation for ∇ . If we assume that the unknown variables are $\dot{\gamma}$ and F , then the equivalent of Theorem 3.1 is:

Proposition 7.1. *If the control input is assumed to be in the class of continuous functions over \mathcal{M} , then the extremals of the cost function \mathcal{J} can be obtained from the solutions of the following system of differential equations in the unknowns $\dot{\gamma}$ and F :*

$$\nabla_{\dot{\gamma}}^2 F + R(F, \dot{\gamma})\dot{\gamma} = 0 \quad \text{subject to } \nabla_{\dot{\gamma}}\dot{\gamma} = F \tag{64}$$

with the boundary conditions $\gamma(t_0) = g_0, \dot{\gamma}(t_0) = v_0, F(t_0) = \nabla_{\dot{\gamma}}\dot{\gamma}|_{\gamma=g_0, \dot{\gamma}=v_0}$ and $\gamma(t_f) = g_f$.

Indeed the solution $\dot{\gamma}$ of the problem is not a constant velocity vector (*i.e.*, the tangent vector of a geodesic curve). The trajectory γ itself, if needed, can be recovered by integration of $\dot{\gamma}$ from the initial condition g_0 .

From (62), instead of the acceleration, the cost functional (18) could be formulated in terms of the input covector forces without any substantial modification: $\tilde{\mathcal{J}} = \frac{1}{2} \int_0^T \langle \tilde{F}, \tilde{F} \rangle dt = \frac{1}{2} \int_0^T \langle \mathbb{I}\nabla_{\dot{\gamma}}\dot{\gamma}, \mathbb{I}\nabla_{\dot{\gamma}}\dot{\gamma} \rangle dt$.

7.3. Reduction of the optimal control problem

If the mechanical system has body fixed actuators, the input vector fields are already left invariant $F = \gamma\mathcal{F}$. Therefore the reduction process under investigation in the previous sections comes as natural simplification also for the control problems. For example, the forced second order vector field (63) reduces to

$$\Gamma + F^v = \gamma(\mathcal{J}; -U(\mathcal{J}, \mathcal{J}) + \mathcal{F}).$$

Assume that the left invariant input distribution \mathcal{F} has the coordinate expression $\mathcal{F} = f^i A_i$.

Proposition 7.2. *The covariant derivatives of the input vector distribution \mathcal{F} are:*

$$\nabla_{\dot{\gamma}}\gamma\mathcal{F} = \nabla^2_{\dot{\gamma}}\dot{\gamma} = \gamma(\dot{\mathcal{F}} + \nabla_{\mathcal{J}}\mathcal{F}) \tag{65}$$

$$\nabla_{\dot{\gamma}}^2\gamma\mathcal{F} = \nabla^3_{\dot{\gamma}}\dot{\gamma} = \gamma(\ddot{\mathcal{F}} + 2\nabla_{\mathcal{J}}\dot{\mathcal{F}} + \nabla_{\dot{\mathcal{J}}}\mathcal{F} + \nabla^2_{\mathcal{J}}\mathcal{F}). \tag{66}$$

Proof. Equation (66) is shown in the same way as (33). For (66), it can be obtained from computations of the same type as in Proposition 6.1.

$$\begin{aligned} \nabla_{\dot{\gamma}}^2\gamma\mathcal{F} &= \nabla_{\dot{\gamma}}\left(\gamma(\dot{\mathcal{F}} + \nabla_{\mathcal{J}}\mathcal{F})\right) \\ &= \mathcal{L}_{\dot{\gamma}}\left(\frac{\partial f^j}{\partial t}\right)\gamma A_j + \mathcal{L}_{\dot{\gamma}}(\alpha^i f^j)\nabla_{A_i}A_j + \alpha^i \frac{\partial f^j}{\partial t}\gamma\nabla_{A_i}A_j + \gamma\alpha^k\alpha^i f^j\nabla_{A_k}\nabla_{A_i}A_j \\ &= \gamma\left(\frac{\partial^2 f^j}{\partial t^2}A_j + \left(\alpha^i \frac{\partial f^j}{\partial t} + \frac{\partial\alpha^i}{\partial t}f^j + \alpha^i \frac{\partial f^j}{\partial t}\right)\nabla_{A_i}A_j + \alpha^k\alpha^i f^j\nabla_{A_k}\nabla_{A_i}A_j\right) \\ &= \gamma(\ddot{\mathcal{F}} + \nabla_{\dot{\mathcal{J}}}\mathcal{F} + 2\nabla_{\mathcal{J}}\dot{\mathcal{F}} + \nabla^2_{\mathcal{J}}\mathcal{F}). \end{aligned} \quad \square$$

Therefore, the expression for the control \mathcal{F} is obtained from (64) and the C^2 dynamic interpolation problem becomes:

Proposition 7.3. *Under the same assumptions of Proposition 7.1, the left invariant control input \mathcal{F} that generates a smooth trajectory which is an extremal of \mathcal{J} is given by the solution of*

$$\ddot{\mathcal{F}} + 2\nabla_{\mathcal{T}}\dot{\mathcal{F}} + \nabla_{\dot{\mathcal{J}}}\mathcal{F} + \nabla^2_{\mathcal{T}}\mathcal{F} + R(\mathcal{F}, \mathcal{T})\mathcal{F} = 0 \quad \text{subject to} \quad \dot{\mathcal{J}} + \nabla_{\mathcal{T}}\mathcal{J} = \mathcal{F} \quad (67)$$

from the boundary conditions $\gamma(t_0) = g_0$, $\gamma(t_f) = g_f$, $V_0 = g_0^{-1}v_0$ and $\mathcal{F}_0 = g_0^{-1}F(t_0)$.

Proof. Since R is a left invariant tensor, this is the straightforward substitution in (64) of equation (66). \square

8. EXAMPLE: $SE(3)$

The *Special Euclidean Group* $SE(3) = SO(3) \otimes \mathbb{R}^3$ is the Lie group of rigid body transformations in \mathbb{R}^3 , i.e., the group of isometric rotations and translations. It is of widespread use in Robotics, see the books [21, 25], and in the study of mechanical control systems, see [1, 4, 5, 16] for a few examples.

Using homogeneous coordinates,

$$SE(3) = \left\{ g \in Gl_4(\mathbb{R}), \quad g = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \text{ s.t. } R \in SO(3) \text{ and } p \in \mathbb{R}^3 \right\}$$

with $SO(3) = \{R \in Gl_3(\mathbb{R}) \mid RR^T = I_3 \text{ and } \det R = +1\}$. The Lie algebra of $SE(3)$ is

$$\mathfrak{se}(3) = \left\{ \mathcal{X} \in M_4(\mathbb{R}), \quad \text{s.t. } \mathcal{X} = \begin{bmatrix} \hat{\omega}_{\mathcal{X}} & v_{\mathcal{X}} \\ 0 & 0 \end{bmatrix} \text{ with } \hat{\omega}_{\mathcal{X}} \in \mathfrak{so}(3) \text{ and } v_{\mathcal{X}} \in \mathbb{R}^3 \right\}$$

with $\mathfrak{so}(3) = \{\hat{\omega}_{\mathcal{X}} \in M_3(\mathbb{R}) \text{ s.t. } \hat{\omega}_{\mathcal{X}}^T = -\hat{\omega}_{\mathcal{X}}\}$ and $\hat{\cdot} : \mathbb{R}^3 \rightarrow \mathfrak{so}(3)$ such that $\hat{\omega}_{\mathcal{X}}\sigma = \omega_{\mathcal{X}} \times \sigma \forall \sigma \in \mathbb{R}^3$.

The Lie group exponential map gives the one-parameter curves corresponding to constant generators in $\mathfrak{se}(3)$ i.e., to the orbits of (complete) constant vector fields and their left/right translations.

For $SO(3)$ and $SE(3)$, the Lie group exponential map corresponds to the ordinary matrix exponential and closed form formulae are available. In $SO(3)$ one can use the so-called Rodriguez' formula:

$$e^{\hat{\omega}_{\mathcal{X}}} = I + \frac{\sin \|\omega_{\mathcal{X}}\|}{\|\omega_{\mathcal{X}}\|} \hat{\omega}_{\mathcal{X}} + \frac{1 - \cos \|\omega_{\mathcal{X}}\|}{\|\omega_{\mathcal{X}}\|^2} \hat{\omega}_{\mathcal{X}}^2$$

while in $SE(3)$:

$$\begin{aligned} e : \quad \mathfrak{se}(3) &\rightarrow SE(3) \\ \mathcal{X} = \begin{bmatrix} \hat{\omega}_{\mathcal{X}} & v_{\mathcal{X}} \\ 0_{3 \times 1} & 0 \end{bmatrix} &\mapsto \begin{bmatrix} e^{\hat{\omega}_{\mathcal{X}}} & A(\hat{\omega}_{\mathcal{X}})v_{\mathcal{X}} \\ 0 & 1 \end{bmatrix} \end{aligned} \quad (68)$$

where

$$A(\hat{\omega}_{\mathcal{X}}) = I + \frac{1 - \cos \|\omega_{\mathcal{X}}\|}{\|\omega_{\mathcal{X}}\|^2} \hat{\omega}_{\mathcal{X}} + \frac{\|\omega_{\mathcal{X}}\| - \sin \|\omega_{\mathcal{X}}\|}{\|\omega_{\mathcal{X}}\|^3} \hat{\omega}_{\mathcal{X}}^2.$$

In $SE(3)$, the exponential map being onto means that every two elements can be connected by a one-parameter curve called *screw*. Its (normalized) constant infinitesimal generator is called *twist* and corresponds to the axis of the rigid body rototranslation.

The derivation of the adjoint map $\text{Ad}_g(\mathcal{Y}) = L_{g*}R_{g^{-1}*}\mathcal{Y} = g\mathcal{Y}g^{-1}$ with respect to $g = e^{t\mathcal{X}}$, $\mathcal{X} \in \mathfrak{se}(3)$, at the identity of the group

$$\text{ad}_{\mathcal{X}} = \left. \frac{d}{dt} (\text{Ad}_{e^{t\mathcal{X}}}) \right|_{t=0}$$

gives the Lie bracket $\text{ad}_X(\mathcal{Y}) = [X, \mathcal{Y}] = X\mathcal{Y} - \mathcal{Y}X$. The linear representations of the operators $\text{Ad}_g(\cdot)$ and $\text{ad}_X(\cdot)$ is:

$$\text{Ad}_g = \begin{bmatrix} R & 0 \\ \hat{p}R & R \end{bmatrix}, \quad \text{ad}_X = \begin{bmatrix} \hat{\omega}_X & 0 \\ \hat{v}_X & \hat{\omega}_X \end{bmatrix}. \tag{69}$$

The natural affine connection that can be associated to a biinvariant nondegenerate symmetric $(0, 2)$ -tensor is called the (0) -connection and is studied by Cartan in [7]. However, since the corresponding quadratic form is nondegenerate but not positive definite, it is not compatible with the standard definition of kinetic energy of a rigid body in G because of the negative energy that can be associated along certain trajectories. Therefore we neglect it and concentrate instead on a positive definite \mathbb{I} . Because of the lack of bi-invariance of \mathbb{I} , its natural connection is not among the ‘‘canonical’’ ones studied in the classical literature [7,12], but rather it can be seen as the torsion-free metric connection of a trivially reductive homogeneous space with respect to the left action on itself and studied accordingly (see for example [23], Sect. 13).

For the metric structure we are adopting, the Riemannian exponential map Exp differs from the Lie group exponential map (68). In fact, disregarding the action $SO(3) \rightarrow \text{End}(\mathbb{R}^3)$ means dropping the $A(\hat{\omega})$ term:

$$\begin{aligned} \text{Exp} : \quad \mathfrak{se}(3) &\rightarrow SE(3) \\ \mathcal{X} = \begin{bmatrix} \hat{\omega}_X & v_X \\ 0_{3 \times 1} & 0 \end{bmatrix} &\mapsto \begin{bmatrix} e^{\hat{\omega}_X} & v_X \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{70}$$

This corresponds to the exponential map for the *direct* product of Lie groups $SO(3) \otimes \mathbb{R}^3$ which pairs the geodesics of $SO(3)$ and the straight lines of \mathbb{R}^3 .

From (69), the expressions for the coadjoint and infinitesimal coadjoint actions $\text{Ad}_{g^{-1}}^*$ and ad_X^* are:

$$\begin{aligned} \text{Ad}_{g^{-1}}^* &= (\text{Ad}_g)^{-T} = \begin{bmatrix} R & \hat{p}R \\ 0 & R \end{bmatrix}, \\ \text{ad}_X^* &= - \frac{d}{dt} \text{Ad}_{e^{-tX}}^* \Big|_{t=0} = -(\text{ad}_X)^T = \begin{bmatrix} -\hat{\omega}_X & -\hat{v}_X \\ 0 & -\hat{\omega}_X \end{bmatrix}. \end{aligned}$$

In (27), when we compute the covariant derivative of \mathcal{Y} along \mathcal{X} , due to the semidirect action of $SO(3)$ on \mathbb{R}^3 , the terms $\text{ad}_X^* \mathbb{I} \mathcal{Y}$ are nonnull, even when the metric tensor is diagonal, $\mathbb{I} = I$. In this case \mathbb{I} can be pulled out and (with abuse of notation)

$$\text{ad}_X^* \mathbb{I} \mathcal{Y} = \text{ad}_X^* \mathcal{Y} = \begin{bmatrix} 0 \\ -\hat{\omega}_X v_Y \end{bmatrix} = \begin{bmatrix} 0 \\ v_Y \times \omega_X \end{bmatrix} \neq 0.$$

In particular then

$$U(\mathcal{X}, \mathcal{Y}) = -\frac{1}{2} (\text{ad}_X^* \mathcal{Y} + \text{ad}_Y^* \mathcal{X}) = \frac{1}{2} \begin{bmatrix} 0 \\ \hat{\omega}_X v_Y + \hat{\omega}_Y v_X \end{bmatrix} \in \mathbb{R}^3. \tag{71}$$

Since $\text{ad}_X \mathcal{Y} = \begin{bmatrix} \hat{\omega}_X \omega_Y \\ \hat{\omega}_X v_Y - \hat{\omega}_Y v_X \end{bmatrix}$, the covariant derivative is

$$\nabla_X \mathcal{Y} = \frac{1}{2} \text{ad}_X \mathcal{Y} + U(\mathcal{X}, \mathcal{Y}) = \begin{bmatrix} \frac{1}{2} \hat{\omega}_X \omega_Y \\ \hat{\omega}_X v_Y \end{bmatrix}. \tag{72}$$

The linear map A_Y of Proposition 2.1 is

$$A_Y = \begin{bmatrix} \frac{1}{2} \hat{\omega}_Y & 0 \\ \hat{v}_Y & 0 \end{bmatrix}$$

which is skew-symmetric with respect to $\langle \cdot, \cdot \rangle$ for all X if and only if $Y \in \mathfrak{so}(3)$. Therefore $Y \notin \mathfrak{so}(3)$ is not an infinitesimal isometry for $\mathbb{I} = I$ *i.e.*, all left invariant Killing vector fields of $(SE(3), \mathbb{I})$ are of the form $Y = \begin{bmatrix} \omega_Y \\ 0 \end{bmatrix}$.

In fact, $\mathcal{Y} \in \mathfrak{so}(3)$ implies that $e^{\mathcal{Y}} = \text{Exp}(\mathcal{Y})$ and therefore one parameter subgroups generated by \mathcal{Y} coincide with geodesics of \mathbb{I} . Notice from (71) that this does not imply $U = 0$, only that for all $\mathcal{X}, \mathcal{Z} \in \mathfrak{se}(3)$ the contribution of $U(\mathcal{X}, \mathcal{Z}) \in \mathbb{R}^3$ along \mathcal{Y} is zero:

$$\langle \text{ad}_{\mathcal{Y}}\mathcal{X}, \mathcal{Z} \rangle + \langle \mathcal{X}, \text{ad}_{\mathcal{Y}}\mathcal{Z} \rangle = 2\langle U(\mathcal{X}, \mathcal{Z}), \mathcal{Y} \rangle = 0.$$

However, this holds true only in virtue of the choice of a diagonal metric tensor.

If $\mathcal{X} = \mathcal{Y}$ one finds the usual Euler equations for rigid bodies (see (42))

$$\dot{\mathcal{X}} = \begin{bmatrix} \dot{\omega}_x \\ \dot{v}_x \end{bmatrix} = -\nabla_x \mathcal{X} = \begin{bmatrix} 0 \\ -\widehat{\omega}_x v_x \end{bmatrix}$$

with geodesic spray

$$\Gamma = \gamma \mathcal{X}^h = (\gamma \mathcal{X}; -\gamma U(\mathcal{X}, \mathcal{X})) = \left(\gamma \begin{bmatrix} \omega_x \\ v_x \end{bmatrix}; \gamma \begin{bmatrix} 0 \\ \widehat{\omega}_x v_x \end{bmatrix} \right).$$

Similarly to (72), we have

$$\nabla_{\dot{\mathcal{X}}} \mathcal{Y} = \begin{bmatrix} \frac{1}{2} \widehat{\omega}_x \omega_y \\ \widehat{\omega}_x v_y \end{bmatrix}, \quad \nabla_x \dot{\mathcal{Y}} = \begin{bmatrix} \frac{1}{2} \widehat{\omega}_x \dot{\omega}_y \\ \widehat{\omega}_x \dot{v}_y \end{bmatrix}, \quad \nabla_x \nabla_y \mathcal{Z} = \frac{1}{4} \begin{bmatrix} \widehat{\omega}_x \widehat{\omega}_y \omega_z \\ 4 \widehat{\omega}_x \widehat{\omega}_y v_z \end{bmatrix} \tag{73}$$

and

$$\nabla_{\mathcal{W}} \nabla_x \nabla_y \mathcal{Z} = \frac{1}{8} \begin{bmatrix} \widehat{\omega}_{\mathcal{W}} \widehat{\omega}_x \widehat{\omega}_y \omega_z \\ 8 \widehat{\omega}_{\mathcal{W}} \widehat{\omega}_x \widehat{\omega}_y v_z \end{bmatrix}.$$

From (10), using some vector algebra with the convention that $a \times b \times c = a \times (b \times c)$, after a few calculations, the values of the curvature tensor are given by²

$$R(\mathcal{X}, \mathcal{Y})\mathcal{Z} = \begin{bmatrix} (\omega_x \times \omega_y) \times \omega_z \\ r_2 \end{bmatrix}$$

where $r_2 \in \mathbb{R}^3$ is

$$\begin{aligned} r_2 = & \frac{3}{4} \omega_x \times \omega_y \times v_z + \frac{3}{4} \omega_y \times \omega_x \times v_z + \frac{1}{4} (\omega_x \times \omega_y) \times v_z \\ & + \frac{1}{4} \omega_x \times \omega_z \times v_y + \frac{3}{4} \omega_z \times \omega_x \times v_y - \frac{1}{4} (\omega_x \times \omega_z) \times v_y \\ & + \frac{1}{4} \omega_y \times \omega_z \times v_x - \frac{3}{4} \omega_z \times \omega_y \times v_x - \frac{1}{4} (\omega_y \times \omega_z) \times v_x. \end{aligned}$$

²Warning: in [26] a wrong expression is reported.

In our case, since $U(\mathcal{X}, \mathcal{Y})$ is null if \mathcal{X}, \mathcal{Y} are both in $\mathfrak{so}(3)$ or in \mathbb{R}^3 , the values of the curvature are given by:

- if $\mathcal{X}, \mathcal{Y} \in \mathfrak{so}(3)$

$$\begin{aligned} R(\mathcal{X}, \mathcal{Y})\mathcal{Z} &= -\frac{1}{4}[[\mathcal{X}, \mathcal{Y}], \mathcal{Z}] \\ &= -\frac{1}{4} \left[\begin{array}{c} (\omega_{\mathcal{X}} \times \omega_{\mathcal{Y}}) \times \omega_{\mathcal{Z}} \\ (\omega_{\mathcal{X}} \times \omega_{\mathcal{Y}}) \times v_{\mathcal{Z}} - \omega_{\mathcal{Z}} \times \omega_{\mathcal{X}} \times v_{\mathcal{Y}} + \omega_{\mathcal{Z}} \times \omega_{\mathcal{Y}} \times v_{\mathcal{X}} \end{array} \right]; \end{aligned}$$

- if $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^3$ $R = 0$;
- if $\mathcal{X} \in \mathfrak{so}(3), \mathcal{Y} \in \mathbb{R}^3$

$$R(\mathcal{X}, \mathcal{Y})\mathcal{Z} = \begin{bmatrix} 0 \\ \frac{1}{4}\omega_{\mathcal{X}} \times \omega_{\mathcal{Z}} \times v_{\mathcal{Y}} + \frac{3}{4}\omega_{\mathcal{Z}} \times \omega_{\mathcal{X}} \times v_{\mathcal{Y}} - \frac{1}{4}(\omega_{\mathcal{X}} \times \omega_{\mathcal{Z}}) \times v_{\mathcal{Y}} \end{bmatrix}.$$

Notice, furthermore, that it is easy to compute the sectional curvatures of $(SE(3), I)$

$$K(\mathcal{X}, \mathcal{Y}) = \frac{\langle R(\mathcal{X}, \mathcal{Y})\mathcal{X}, \mathcal{Y} \rangle}{|\mathcal{X}|^2|\mathcal{Y}|^2 - \langle \mathcal{X}, \mathcal{Y} \rangle^2}.$$

In fact, from [2,3], the general expression for the (non-normalized) two-plane curvature for a semidirect product is given by:

$$\begin{aligned} \langle R(\mathcal{X}, \mathcal{Y})\mathcal{X}, \mathcal{Y} \rangle &= -\frac{3}{4}\langle [\mathcal{X}, \mathcal{Y}], [\mathcal{X}, \mathcal{Y}] \rangle - \frac{1}{2}\langle [\mathcal{X}, [\mathcal{X}, \mathcal{Y}]], \mathcal{Y} \rangle - \frac{1}{2}\langle [\mathcal{Y}, [\mathcal{Y}, \mathcal{X}]], \mathcal{X} \rangle \\ &\quad + \langle U(\mathcal{X}, \mathcal{Y}), U(\mathcal{X}, \mathcal{Y}) \rangle - \langle U(\mathcal{X}, \mathcal{X}), U(\mathcal{Y}, \mathcal{Y}) \rangle \end{aligned}$$

therefore

- if $\mathcal{X}, \mathcal{Y} \in \mathfrak{so}(3)$

$$\langle R(\mathcal{X}, \mathcal{Y})\mathcal{X}, \mathcal{Y} \rangle = \frac{1}{4}\langle [\mathcal{X}, \mathcal{Y}], [\mathcal{X}, \mathcal{Y}] \rangle = \frac{1}{4}|\omega_{\mathcal{X}} \times \omega_{\mathcal{Y}}|^2;$$

- if $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^3$ $R \equiv 0$;
- if $\mathcal{X} \in \mathfrak{so}(3), \mathcal{Y} \in \mathbb{R}^3$ then $\langle [\mathcal{X}, [\mathcal{X}, \mathcal{Y}]], \mathcal{Y} \rangle = \langle [\mathcal{Y}, [\mathcal{Y}, \mathcal{X}]], \mathcal{X} \rangle = 0$ and

$$\langle R(\mathcal{X}, \mathcal{Y})\mathcal{X}, \mathcal{Y} \rangle = -\frac{3}{4}\langle \text{ad}_{\mathcal{X}}\mathcal{Y}, \text{ad}_{\mathcal{X}}\mathcal{Y} \rangle + \langle U(\mathcal{X}, \mathcal{Y}), U(\mathcal{X}, \mathcal{Y}) \rangle = -\frac{1}{2}|\omega_{\mathcal{X}} \times v_{\mathcal{Y}}|^2.$$

If we consider the curve γ of \mathfrak{g} -valued tangent vector field $\mathcal{T} = \begin{bmatrix} \omega_{\mathcal{T}} \\ v_{\mathcal{T}} \end{bmatrix}$, then the necessary condition (61) corresponds to the system:

$$\begin{aligned} \ddot{\omega}_{\mathcal{T}} + \frac{3}{2}\omega_{\mathcal{T}} \times \ddot{\omega}_{\mathcal{T}} + \frac{1}{2}\omega_{\mathcal{T}} \times \omega_{\mathcal{T}} \times \dot{\omega}_{\mathcal{T}} + \frac{1}{2}\omega_{\mathcal{T}} \times \dot{\omega}_{\mathcal{T}} \times \omega_{\mathcal{T}} + (\dot{\omega}_{\mathcal{T}} \times \omega_{\mathcal{T}}) \times \omega_{\mathcal{T}} &= 0 \\ \ddot{v}_{\mathcal{T}} + 3\omega_{\mathcal{T}} \times \ddot{v}_{\mathcal{T}} + 3\dot{\omega}_{\mathcal{T}} \times \dot{v}_{\mathcal{T}} + \ddot{\omega}_{\mathcal{T}} \times v_{\mathcal{T}} + \frac{5}{2}\omega_{\mathcal{T}} \times \omega_{\mathcal{T}} \times \dot{v}_{\mathcal{T}} + \\ + \frac{7}{2}\omega_{\mathcal{T}} \times \dot{\omega}_{\mathcal{T}} \times v_{\mathcal{T}} + 2\dot{\omega}_{\mathcal{T}} \times \omega_{\mathcal{T}} \times v_{\mathcal{T}} + \frac{1}{2}\omega_{\mathcal{T}} \times \omega_{\mathcal{T}} \times \omega_{\mathcal{T}} \times v_{\mathcal{T}} &= 0. \end{aligned}$$

Substituting $\dot{\mathcal{T}} + \nabla_{\mathcal{T}}\mathcal{T}$ with the control input $\mathcal{F} = \begin{bmatrix} \omega_{\mathcal{F}} \\ v_{\mathcal{F}} \end{bmatrix}$ of the mechanical system (62), equation (67) becomes the system of ordinary differential equations which are linear and second order in \mathcal{F} , quadratic and first order

in \mathcal{T} :

$$\begin{aligned} \ddot{\omega}_{\mathcal{F}} + \omega_{\mathcal{F}} \times \dot{\omega}_{\mathcal{F}} + \frac{1}{2} \dot{\omega}_{\mathcal{F}} \times \omega_{\mathcal{F}} + \frac{1}{4} \omega_{\mathcal{F}} \times \omega_{\mathcal{F}} \times \omega_{\mathcal{F}} + (\omega_{\mathcal{F}} \times \omega_{\mathcal{F}}) \times \omega_{\mathcal{F}} &= 0 \\ \ddot{v}_{\mathcal{F}} + 2\omega_{\mathcal{F}} \times \dot{v}_{\mathcal{F}} + \dot{\omega}_{\mathcal{F}} \times v_{\mathcal{F}} - \frac{1}{4} \omega_{\mathcal{F}} \times \omega_{\mathcal{F}} \times v_{\mathcal{F}} + \\ &+ \omega_{\mathcal{F}} \times \omega_{\mathcal{F}} \times v_{\mathcal{F}} + \frac{3}{2} \omega_{\mathcal{F}} \times \omega_{\mathcal{F}} \times v_{\mathcal{F}} = 0. \end{aligned}$$

Acknowledgements. The author wishes to thank one of the reviewers for several helpful comments.

REFERENCES

- [1] C. Altafini, Geometric motion control for a kinematically redundant robotic chain: Application to a holonomic mobile manipulator. *J. Rob. Syst.* **20** (2003) 211-227.
- [2] V.I. Arnold, Math. methods of Classical Mechanics. 2nd ed., *Grad. Texts Math.* **60** (1989).
- [3] L. Berard-Bergery, Sur la courbure des métriques Riemanniennes invariantes des groupes de Lie et des espaces homogènes. *Ann. Sci. Ecole National Superior* **11** (1978) 543.
- [4] F. Bullo, N. Leonard and A. Lewis, Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups. *IEEE Trans. Autom. Control* **45** (2000) 1437-1454.
- [5] F. Bullo and R. Murray, Tracking for fully actuated mechanical systems: a geometric framework. *Automatica* **35** (1999) 17-34.
- [6] M. Camarinha, F. Silva Leite and P. Crouch, Second order optimality conditions for an higher order variational problem on a Riemannian manifold, in *Proc. 35th Conf. on Decision and Control*. Kobe, Japan, December (1996) 1636-1641.
- [7] E. Cartan, La géométrie des groupes de transformations, in *Œuvres complètes* **2**, part I. Gauthier-Villars, Paris, France (1953) 673-792.
- [8] H. Cendra, D. Holm, J. Marsden and T. Ratiu, Lagrangian reduction, the Euler-Poincaré equations and semidirect products. *Amer. Math. Soc. Transl.* **186** (1998) 1-25.
- [9] M. Crampin and F. Pirani, *Applicable differential geometry*. London Mathematical Society Lecture notes. Cambridge University Press, Cambridge, UK (1986).
- [10] P.E. Crouch and F. Silva Leite, The dynamic interpolation problem on Riemannian manifolds, Lie groups and symmetric spaces. *J. Dynam. Control Syst.* **1** (1995) 177-202.
- [11] M. do Carmo, *Riemannian geometry*. Birkhäuser, Boston (1992).
- [12] L. Eisenhart, *Riemannian geometry*. Princeton University Press, Princeton (1966).
- [13] V. Jurdjevic, *Geometric Control Theory*. *Cambridge Stud. Adv. Math.* Cambridge University Press, Cambridge, UK (1996).
- [14] S. Kobayashi and K. Nomizu, *Foundations of differential geometry I and II*. Interscience Publisher, New York (1963) and (1969).
- [15] J. Lee, *Riemannian manifolds. An introduction to curvature*. Springer, New York, NY (1997).
- [16] A. Lewis and R. Murray, Configuration controllability of simple mechanical control systems. *SIAM J. Control Optim.* **35** (1997) 766-790.
- [17] A. Lewis and R. Murray, Decompositions for control systems on manifolds with an affine connection. *Syst. Control Lett.* **31** (1997) 199-205.
- [18] J. Marsden, *Lectures on Mechanics*. Cambridge University Press, Cambridge (1992).
- [19] J. Marsden and T. Ratiu, Introduction to mechanics and symmetry, Springer-Verlag, 2nd ed., *Texts Appl. Math.* **17** (1999).
- [20] J. Milnor, Curvature of left invariant metrics on Lie groups. *Adv. Math.* **21** (1976) 293-329.
- [21] R. Murray, Z. Li and S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. CRC Press (1994).
- [22] L. Noakes, G. Heinzinger and B. Paden, Cubic splines on curved spaces. *IMA J. Math. Control Inform.* **12** (1989) 465-473.
- [23] K. Nomizu, Invariant affine connections on homogeneous spaces. *Amer. J. Math.* **76** (1954) 33-65.
- [24] F. Park and B. Ravani, Bézier curves on Riemannian manifolds and Lie groups with kinematic applications. *ASME J. Mech. design* **117** (1995) 36-40.
- [25] J. M. Selig, *Geometrical methods in Robotics*. Springer, New York, NY (1996).
- [26] M. Zefran, V. Kumar and C. Croke, On the generation of smooth three-dimensional rigid body motions. *IEEE Trans. Robot. Automat.* **14** (1998) 576-589.