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

CLAUDIO ALTAFINI†

Abstract. For an invariant Lagrangian equal to kinetic energy and defined on a semidirect product of Lie groups, the variational problems can be reduced using the group symmetry. Choosing the Riemannian 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 due to 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 contain more terms than the original ones. The reduced Euler-Lagrange equations are well-known under the name of Euler-Poincaré equations. We treat in a similar way the reduction of second order variational problems corresponding to geometric splines on the Lie group. Here the problems connected with the semidirect structure are emphasized and a number of extra terms is appearing in the reduction. If the Lagrangian corresponds to a fully actuated mechanical system, then the resulting necessary condition can be expressed directly in terms of the control input. As an application, the case of a rigid body on the Special Euclidean group is considered.

Key words. Lie group, semidirect product, second order variational problems, reduction, group symmetry, optimal control

AMS subject classifications. 22F30, 70Q05, 93B29, 49J15, 70H30, 70H33

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 [16]. 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 constituted 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 therefore the Riemannian connection induced by $\mathbb{I}$ is (in the language of [11], Ch.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

*This work was supported by the Swedish Foundation for Strategic Research through the Center for Autonomous Systems at KTH and by a grant from the "C.M. Lerici" Foundation

†Division of Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, SE-10044 Stockholm, Sweden; (altafini@math.kth.se).

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 [2, 3]). 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 $\mathbb{I}$.

The reduction process can be seen as the projection map $\pi : TG \to \mathfrak{g}$ of a globally trivial principal fiber bundle with base manifold $\mathfrak{g}$ and structure group G. For matrix groups, such a construction resembles closely a G-structure obtained from the frame bundle i.e. the collection of all the linear changes of basis on the tangent bundle, but in general it has to be intended as induced by left (or right) invariance of G. The mismatch between Lie group exponential map and Riemannian exponential map implies that the horizontal vectors determined by the Riemannian 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.

The motivation behind this work is generating 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 $SE(3)$. The presence of control inputs allows 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. 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.

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 [8, 18] 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 $\mathbb{R}^n$. While the reduction for compact Lie groups is quite straightforward (see [8]), in semidirect products of Lie groups like $SE(3)$ the extra difficulties mentioned above all arise. In particular, for the same reason that a drift term appears in the Euler-Poincaré equations, several extra components arise in the reduced necessary conditions for optimality of the new cost functional. Their explicit calculation is the main contribution of this paper.

2. Mathematical preliminaries. A *Riemannian metric* on a smooth 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 $T_x\mathcal{M}$, $\langle X, Y \rangle = \mathbb{I}(X, Y)$, for $X, Y \in T_x\mathcal{M}$. One important property of Riemannian metrics is that they allow to convert vectors to covectors and viceversa. In particular, at each $x \in \mathcal{M}$ this allows to view the metric tensor as a map $\mathbb{I} : T_x\mathcal{M} \to T_x^*\mathcal{M}$.

Call $\mathcal{D}(\mathcal{M})$ the space of smooth sections of $T\mathcal{M}$. Elements of $\mathcal{D}(\mathcal{M})$ 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 bilinear in X and Y
2. $\nabla_{fX} Y = f\nabla_X Y$
3. $\nabla_X (fY) = f\nabla_X Y + \ (\mathcal{L}_X)Y$

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 $x^1, \dots, x^n$, the covariant derivative is

$$(\nabla_X Y)^k = \frac{\partial Y^k}{\partial x^i} X^i + \Gamma^k_{ij} X^i Y^j \tag{2.1}$$

where $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^i \frac{\partial}{\partial x^i}$, $\nabla_X Y = (\nabla_X Y)^k \frac{\partial}{\partial x^k}$ and the n^3 quantities Γ^k_{ij} are called *Christoffel symbols* and are given in by $\nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j}\right) = \Gamma^k_{ij} \frac{\partial}{\partial x^k}$. 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 and in fact it reduces to the standard notion of Euclidean acceleration if $\mathcal{M} = \mathbb{R}^n$ and we choose the so-called the Euclidean connection $\nabla_X Y = XY^k \frac{\partial}{\partial x^k} = \left(X^i \frac{\partial Y^k}{\partial x^i}\right) \frac{\partial}{\partial x^k}$, i.e. the vector field whose components are the directional derivatives of the components of Y along X. The length of the smooth curve γ is measured by the functional

$$\ell(\gamma) = \int_{t_0}^{t_f} \langle \dot\gamma(t), \dot\gamma(t)\rangle^{\frac{1}{2}} dt \tag{2.2}$$

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*:

$$\frac{D}{dt}\left(\frac{d\gamma}{dt}\right) = \nabla_{\dot\gamma(t)} \dot\gamma(t) = 0 \tag{2.3}$$

Geodesic motion corresponds to constant velocity and it gives an extremum of the length functional (2.2), 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 $\frac{dY^k}{dt} + \Gamma^k_{ij} \dot x^i Y^j = 0$ and the one for geodesic motion

$$\ddot x^k + \Gamma^k_{ij} \dot x^i \dot x^j = 0 \tag{2.4}$$

Along $\gamma : (t_0, t_f) \to \mathcal{M}$, for $t_0 \le t_1 \le t_2 \le t_f$, parallel transport defines an operator

$$\begin{aligned} P_{(t_1,t_2)} \ &: T_{\gamma(t_1)}\mathcal{M} \to T_{\gamma(t_2)}\mathcal{M} \\ &X_1 \mapsto X_2 = P_{(t_1,t_2)} X_1 \end{aligned} \tag{2.5}$$

which is a linear isomorphism between tangent spaces.

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

1. ∇ is torsion free:

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{2.6}$$

2. the parallel transport is an isometry

$$Z\langle X, Y\rangle = \langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle \tag{2.7}$$

for all $X, Y, Z \in \mathcal{D}(\mathcal{M})$. Such a connection is called the *Levi-Civita* or *Riemannian connection*. From (2.7), we get the Koszul formula:

$$\langle Z, \nabla_X Y\rangle = \frac{1}{2}\left(Y\langle X, Z\rangle + X\langle Z, Y\rangle - Z\langle X, Y\rangle - \langle [Y, Z], X\rangle + \right. \\ \left. + \langle [Z, X], Y\rangle + \langle [X, Y], Z\rangle\right) \tag{2.8}$$

Condition 2. alone means that ∇ is a metric connection (i.e. $\nabla\mathbb{I} = 0$). The "measure" of the failure of the second covariant derivative to commute is expressed geometrically by the notion of *curvature*, i.e. the map $R : \mathcal{D}(\mathcal{M}) \times \mathcal{D}(\mathcal{M}) \times \mathcal{D}(\mathcal{M}) \to \mathcal{D}(\mathcal{M})$ defined by

$$R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X, Y]}Z \tag{2.9}$$

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

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

2.1. The variational principle of Hamilton. The geodesic equation (2.3) can be obtained from standard calculus of variation on the Riemannian manifold $(\mathcal{M}, \mathbb{I})$, see for example [9]. Given the curve $\gamma : [t_0, t_f] \to \mathcal{M}$ consider *proper variations* of γ i.e. the family of fixed end-point curves $\mathcal{G} : (-\epsilon, \epsilon) \times [t_0, t_f] \to \mathcal{M}$ such that

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

and

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

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 $V(t)$ based at $\gamma(t)$ will be of the type $\mathcal{G}(s, t) = \operatorname{Exp}(sV(t))$. In fact, for a fixed $\bar{t} \in [t_0, t_f]$, if we have $\mathcal{G}^{(\bar{t})}(s) = \operatorname{Exp}(sV(\bar{t}))$ then

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

Another standard result is the *symmetry lemma*, that allows to exchange the order of the mixed second order derivatives along main and transverse curves. Calling

$$S(s, t) = \frac{d}{ds}\mathcal{G}^{(t)}(s) \quad \text{and } T(s, t) = \frac{d}{dt}\mathcal{G}_s(t) \tag{2.10}$$

(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 (2.6), that the vector fields T and S commute $[T,\, S] = 0$. Furthermore, since the Riemannian connection is an isometry, from (2.7) we have

$$\frac{d}{dt}\langle S,\, T\rangle = T\langle S,\, T\rangle = \langle \nabla_T S,\, T\rangle + \langle S,\, \nabla_T T\rangle \tag{2.11}$$

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

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

we have

$$\begin{aligned}
\frac{d}{ds}\mathcal{E}\left(\mathcal{G}_s(t)\right)\Big|_{s=0} &= \frac{d}{ds}\int_{t_0}^{t_f} \langle T(s,\, t),\, T(s,\, t)\rangle dt\Big|_{s=0} \\
&= \int_{t_0}^{t_f} \langle \nabla_S T,\, T\rangle dt\Big|_{s=0} \\
&= \int_{t_0}^{t_f} \langle \nabla_{\delta\gamma}\dot{\gamma},\, \dot{\gamma}\rangle dt = \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}}\delta\gamma,\, \dot{\gamma}\rangle dt \qquad \text{by the symmetry lemma} \\
&= \int_{t_0}^{t_f} \left(\frac{d}{dt}\langle \delta\gamma,\, \dot{\gamma}\rangle - \langle \nabla_{\dot{\gamma}}\dot{\gamma},\, \delta\gamma\rangle\right) dt \qquad \text{by (2.11)} \\
&= \langle \delta\gamma,\, \dot{\gamma}\rangle\big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}}\dot{\gamma},\, \delta\gamma\rangle dt
\end{aligned}$$

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

$$\frac{d}{ds}\mathcal{E}\left(\mathcal{G}_s(t)\right)\Big|_{s=0} = 0 \Longleftrightarrow \nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = 0 \tag{2.12}$$

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

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*

$$\nabla^2_{\dot{\gamma}} V + R(V,\, \dot{\gamma})\dot{\gamma} = 0 \tag{2.13}$$

A vector field is a Jacobi field if and only if it is the variation field of some variations of γ. 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. This implies by (2.5) that the value of $V(t_f) = P_{(t_0,t_f)} V(t_0)$ is uniquely defined from the triple of initial data.

2.2. Second order structures on a Riemannian manifold. Assume that the coordinate chart $x^1, \ldots, x^n$ is valid in a neighborhood U of $x \in \mathcal{M}$. If $v \in T_v\mathcal{M}$ is a tangent vector, its coordinates description is naturally given by $v = v^i \frac{\partial}{\partial x^i}$. If τ :

$T\mathcal{M} \to \mathcal{M}$ is the tangent bundle projection, $(x^1, \ldots, x^n, v^1, \ldots, v^n)$ are called induced coordinates on $\tau^{-1}(U)$ and they provide a basis of tangent vectors of $T_{(x,v)}T\mathcal{M}$: $\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}, \frac{\partial}{\partial v^1}, \ldots, \frac{\partial}{\partial v^n}\right)$. By taking the tangent map τ_* of the projection τ at the point (x, v) of $T\mathcal{M}$, $\tau_* : T_{(x,v)}T\mathcal{M} \to T_{x=\tau(v)}\mathcal{M} = \tau^{-1}(x)$, one can define the *vertical subspace* of the tangent bundle at (x, v)

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

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^h of u from $T_{\tau(v)}\mathcal{M}$ to $T_{(x,v)}T\mathcal{M}$ is

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

So, for example, $\left(\frac{\partial}{\partial x^i}\right)^v = \left(\frac{\partial}{\partial v^i}\right)$ and a basis for $\mathcal{V}_{(x,v)}$ is given by $\left(0, \frac{\partial}{\partial v^i}\right)$.

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 Riemannian 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 [7] Ch. 13). In fact, calling $\sigma^h = (\sigma, V)$ the horizontal lift of the curve σ through (x, v), the condition $\nabla_u V = 0$ (in coordinates $\dot V^i + \Gamma^i_{jk} V^j u^k = 0$) 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 $t = 0$:

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

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

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

or

$$u^h = u^k \frac{\partial}{\partial x^k} - \Gamma^i_{jk} v^j u^k \frac{\partial}{\partial v^i} \tag{2.14}$$

Since $\tau_*(u^h) = u$, horizontal lifts are indeed complementary to the vertical subspace and in this sense they form the *horizontal subspace* $\mathcal{H}_{(x,v)}$ of $T_{(x,v)}T\mathcal{M}$ whose basis is given by the lifting of the $\frac{\partial}{\partial x^k}$:

$$\left(\frac{\partial}{\partial x^k}\right)^h = \frac{\partial}{\partial x^k} - \Gamma^i_{jk} v^j \frac{\partial}{\partial v^i}$$

Equivalently, $\mathcal{H}_{(x,v)}$ can be defined in terms of sections of the tangent bundle, i.e. of smooth maps $\varsigma : \mathcal{M} \to T\mathcal{M}$ such that $\tau(\varsigma(x)) = x \ \forall x \in \mathcal{M}$, by taking the push forward at $u \in T_{\tau(v)}\mathcal{M}$ of the sections that are parallel transported along the vector u

$$\mathcal{H}_{(x,v)} = \{\varsigma_* u \quad \text{s.t.} \ \nabla_u \varsigma = 0\} \tag{2.15}$$

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 (2.14), by using the same coordinate notation as above for $\gamma(t)$

$$\Gamma_{(x,v)} = v^k \frac{\partial}{\partial x^k} - \Gamma^i_{jk} v^j v^k \frac{\partial}{\partial v^i} \tag{2.16}$$

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 (2.4))

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

From (2.16), both components are homogeneous of degree one in the fiber coordinate v^i.

2.3. Simple mechanical control systems. If we add a forcing term to the geodesic equations (2.3), we obtain a so-called *simple mechanical control system* [13] (without potential):

$$\nabla_{\dot{\gamma}}\dot{\gamma} = F(\gamma) \tag{2.17}$$

where $F = (F_1, \ldots, F_n)$ 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 $F_1, \ldots, F_n$ to be linearly independent on $\mathcal{M}$, then we have a *fully actuated mechanical system.* The system of first order differential equations corresponding to (2.17) was shown in [14] 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:

$$\Gamma + F^v \tag{2.18}$$

having integral curves

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

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

3. A second order variational problem. Following [18, 4, 20], 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:

$$\mathcal{J} = \int_{t_0}^{t_f} \langle \nabla_{\dot{\gamma}}\dot{\gamma},\, \nabla_{\dot{\gamma}}\dot{\gamma} \rangle \, dt \tag{3.1}$$

$\mathcal{J}$ 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. ([8, 18]) *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*

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}\nabla_{\dot\gamma}\dot\gamma + R(\nabla_{\dot\gamma}\dot\gamma,\, \dot\gamma)\dot\gamma = 0 \tag{3.2}$$

Proof. The proof has already appeared in the above mentioned references. It is repeated here only for 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:

$$\langle R(V,\, W)Z,\, U\rangle = \langle R(U,\, Z)W,\, V\rangle \tag{3.3}$$

$$\begin{aligned}
\frac{d}{ds}\mathcal{J}\left(\mathcal{G}_s(t)\right)\Big|_{s=0} &= \int_{t_0}^{t_f}\langle\nabla_S\nabla_T T,\, \nabla_T T\rangle dt\Big|_{s=0} \\
&= \int_{t_0}^{t_f}\langle\nabla_T\nabla_S T + R(S,\, T)T,\, \nabla_T T\rangle dt\Big|_{s=0} && \text{by (2.9)} \\
&= \int_{t_0}^{t_f}\langle\nabla_{\dot\gamma}\nabla_{\delta\gamma}\dot\gamma + R(\delta\gamma,\, \dot\gamma)\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle dt \\
&= \int_{t_0}^{t_f}\left(\frac{d}{dt}\langle\nabla_{\delta\gamma}\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle - \langle\nabla^2_{\dot\gamma}\dot\gamma,\, \nabla_{\delta\gamma}\dot\gamma\rangle + \langle R(\delta\gamma, \dot\gamma)\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle\right) dt && \text{by (2.11)} \\
&= \langle\nabla_{\delta\gamma}\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle\big|_{t_0}^{t_f} + \\
&\qquad + \int_{t_0}^{t_f}\left(-\frac{d}{dt}\langle\nabla^2_{\dot\gamma}\dot\gamma,\, \delta\gamma\rangle + \langle\nabla^3_{\dot\gamma}\dot\gamma,\, \delta\gamma\rangle + \langle R(\delta\gamma, \dot\gamma)\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle\right) dt && \text{by (2.11)} \\
&= \left(\langle\nabla_{\delta\gamma}\dot\gamma,\, \nabla_{\dot\gamma}\dot\gamma\rangle - \langle\nabla^2_{\dot\gamma}\dot\gamma,\, \delta\gamma\rangle\right)\big|_{t_0}^{t_f} + \\
&\qquad + \int_{t_0}^{t_f}\left(\langle\nabla^3_{\dot\gamma}\dot\gamma,\, \delta\gamma\rangle + \langle R(\nabla_{\dot\gamma}\dot\gamma, \dot\gamma)\dot\gamma,\, \delta\gamma\rangle\right) dt && \text{by (3.3)}
\end{aligned}$$

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

The condition (3.2) replaces the geodesic condition (2.3) 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. Eq. (2.17) provides the expression for the control corresponding to a solution of (3.2). The full actuation of the mechanical control systems is a sufficient condition for free feasibility of the trajectories of (3.2).

In the extra smoothness assumption that also the control input is continuous between different intervals, we can obtain directly an expression for the controller out of (3.2). In order to set up the problem correctly, one needs an initial value for F i.e.

$F(t_0) = \nabla_{v_0} v_0|_{\gamma=g_0}$ which can replace the final data v_f. This type of problems can be referred to as "C^2 dynamical interpolation problem" (see [8]). If $F(t_0)$ is used in conjunction with v_f and $F(t_f) = \nabla_{v_f} v_f\big|_{\gamma=g_f}$ then one can obtain a C^3 curve in the patching of intervals. Using (2.17) to insert the control input F into (3.2), 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 3.2. *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 by the solutions of the following system of differential equations in the unknowns $\dot\gamma$ and F :*

$$\nabla^2_{\dot\gamma} F + R(F, \dot\gamma)\dot\gamma = 0 \qquad \textit{subject to } \nabla_{\dot\gamma}\dot\gamma = F \tag{3.4}$$

with the boundary conditions $\gamma(t_0) = g_0$, $\dot\gamma(t_0) = v_0$, $F(t_0) = \nabla_{v_0} v_0|_{\gamma=g_0}$ 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 (2.17), instead of the acceleration, the cost functional (3.1) could be formulated in terms of the input covector forces without any substantial modification: $\tilde{\mathcal{J}} = \int_0^T \langle \tilde F, \tilde F\rangle\, dt = \int_0^T \langle \mathbb{I}\nabla_{\dot\gamma}\dot\gamma, \mathbb{I}\nabla_{\dot\gamma}\dot\gamma\rangle\, dt$.

4. Riemannian connection on a semidirect product of Lie groups. For any $g \in G$, a *left translation on* G is a transitive and free action of the group on itself

$$\begin{aligned} L_g : G &\to G \\ h &\mapsto L_g(h) = gh \qquad h \in G \end{aligned} \tag{4.1}$$

Since G is a Lie Group, L_g is a diffeomorphism of G for each g with respect to the identity element e of the group:

$$\begin{aligned} L_g : G &\to G \\ e &\mapsto L_g(e) = g \end{aligned} \tag{4.2}$$

In fact $L_g \circ L_h = gh \Rightarrow (L_g)^{-1} = L_{g^{-1}}$. Similar things hold for a right translation R_g. By deriving the left translation (4.1), we obtain left invariant vector fields. A vector field $\mathcal{X}$ on G is called left invariant if for every $g \in G$ we have $L_g^* \mathcal{X} = \mathcal{X}$ i.e. $(T_h L_g)\, \mathcal{X}(h) = \mathcal{X}(gh)\ \forall\ h \in G$. The set $\mathfrak{X}_L(G)$ of left invariant vectors on G is isomorphic to the Lie algebra $\mathfrak{g} = T_e G$.

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. We are interested in the case in which the metric tensor $\mathbb{I}$ is compatible with the kinetic 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 kinematic 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: $G = K \circledS V$. 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, $K \to \mathrm{Aut}(V)$, so that the group multiplication looks like

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

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

$$[(\mathcal{K}_1, v_1), (\mathcal{K}_2, v_2)] = ([\mathcal{K}_1, \mathcal{K}_2], \mathcal{K}_1 v_2 - \mathcal{K}_2 v_1) \qquad \mathcal{K}_1, \mathcal{K}_2 \in \mathfrak{k}, \ \ v_1, v_2 \in V$$

If K has no nontrivial ideals, V being abelian forms an ideal in $\mathfrak{g}$ 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.

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

$$(k, u_1)u_2 = ku_2 + u_1 \qquad \forall\ (k, u_1) \in G \ \ u_2 \in V$$

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

The Riemannian 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 the elements of an orthonormal basis of left invariant vector fields:

$$\nabla_{gA_i}(gA_j) = g\nabla_{A_i}A_j = \Gamma_{ij}^k g A_k \tag{4.3}$$

for all $g \in G$. From (2.1), 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, Ch.VI of [11]): $(\phi_X)_*(\nabla_Y Z) = \nabla_{(\phi_X)_*(Y)}(\phi_X)_*(Z)\ \forall\ Y, Z \in \mathcal{D}(G)$. The infinitesimal generator above $\mathcal{Y}_G(u)$ is an example of how one-parameter subgroups affinely generated emerge in a semidirect product.

In correspondence of left invariant vector fields $\mathcal{Y}$ and $\mathcal{Z}$, the equation (2.7) for the parallel transport simplifies to

$$\langle \nabla_X \mathcal{Y}, \mathcal{Z} \rangle + \langle \mathcal{Y}, \nabla_X \mathcal{Z} \rangle = 0 \tag{4.4}$$

since $\langle \mathcal{Y}, \mathcal{Z} \rangle$ is a constant.

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

$$\langle \mathrm{ad}_{\mathcal{X}}\mathcal{Z}, \mathcal{Y} \rangle = (\mathrm{ad}_{\mathcal{X}}\mathcal{Z}; \mathbb{I}\mathcal{Y}) = (\mathcal{Z}; \mathrm{ad}_{\mathcal{X}}^*\mathbb{I}\mathcal{Y}) = \langle \mathcal{Z}, \mathbb{I}^{-1}\mathrm{ad}_{\mathcal{X}}^*\mathbb{I}\mathcal{Y} \rangle \tag{4.5}$$

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 Riemannian connection (but *not* of the Lie group exponential map in the case we are considering, see below) in a neighborhood of each point consists of isometries. We have the following equivalent characterizations:

Proposition 4.1. **([11] Prop. 3.2 Ch. VI)** *Given a vector field X on a Riemannian manifold with metric connection $(G, \mathbb{I})$, the following are equivalent:*

(i) *X is a Killing vector field*
(ii) $\mathcal{L}_X \mathbb{I} = 0$
(iii) *the Killing equation holds:*

$$X\langle Y,\, Z\rangle = \langle [X,\, Y],\, Z\rangle + \langle Y,\, [X,\, Z]\rangle \qquad \forall\, Y,\, Z \in\, \mathcal{D}(G) \tag{4.6}$$

or equivalently

$$\langle \nabla_Y X,\, Z\rangle + \langle Y,\, \nabla_Z X\rangle = 0 \qquad \forall\ Y,\, Z\, \in\, \mathcal{D}(G) \tag{4.7}$$

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

From (4.4), if $X \in \mathcal{D}(G)$, for left invariant vector fields $\mathcal{Y}$, $\mathcal{Z} \in \mathfrak{g}$ (4.6) reduces to

$$X\langle \mathcal{Y},\, \mathcal{Z}\rangle = \langle [X,\, \mathcal{Y}],\, \mathcal{Z}\rangle + \langle \mathcal{Y},\, [X,\, \mathcal{Z}]\rangle = 0$$

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 [11]:

THEOREM 4.2. *Given $(G,\, \mathbb{I})$, the Riemannian connection for $\mathbb{I}$ is expressed as*

$$\nabla_{\mathcal{X}} \mathcal{Y} = \frac{1}{2}[\mathcal{X},\, \mathcal{Y}] + U(\mathcal{X},\, \mathcal{Y}) \tag{4.8}$$

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

$$\langle U(\mathcal{X},\, \mathcal{Y}),\, \mathcal{Z}\rangle = \frac{1}{2}\left(\langle \mathrm{ad}_{\mathcal{Z}}\mathcal{X},\, \mathcal{Y}\rangle + \langle \mathcal{X},\, \mathrm{ad}_{\mathcal{Z}}\mathcal{Y}\rangle\right) \tag{4.9}$$

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

COROLLARY 4.3. *If K is semisimple compact and the biinvariant metric is chosen on it, then $U(\mathcal{X},\, \mathcal{Y}) \in V$ for all $\mathcal{X}$, $\mathcal{Y} \in \mathfrak{g}$.*

Applying the pairing (4.5) to $U(\mathcal{X},\, \mathcal{Y})$ then we have:

PROPOSITION 4.4. *The left-invariant covariant derivative* (4.3) *can be expressed as*

$$\nabla_{\mathcal{X}} \mathcal{Y} = \frac{1}{2}\left([\mathcal{X}, \mathcal{Y}] - \mathbb{I}^{-1}\left(\mathrm{ad}^*_{\mathcal{X}}\mathbb{I}\mathcal{Y} + \mathrm{ad}^*_{\mathcal{Y}}\mathbb{I}\mathcal{X}\right)\right) \tag{4.10}$$

Proof. Using (4.5) to extract $\mathcal{Z}$ on both terms on the right hand side expression (4.9) 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 4.5. *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 $\nabla_{\mathcal{X}}\mathcal{X} = 0$ for all the Lie algebra valued vectors $\mathcal{X}$ ([11] Prop.2.9 Ch.X). □

In fact, using the language of [11], $\nabla_{\mathcal{X}}\mathcal{X} = 0$ 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 4.6. *Given a positive definite Riemannian metric on a group manifold G:*

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

Proof. The proof can be obtained from part 2 of Theorem 3.3 Ch.X in [11]. To see it directly use (4.8): assuming the right hand side expression holds true, if $\nabla_{\mathcal{X}}\mathcal{X} \neq 0$ 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 $\nabla_{\mathcal{X}}\mathcal{X} = 0$. 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. □

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 a compact Lie groups with bi-invariant metric: its covariant derivative is well-known to be simply $\nabla_{\mathcal{X}}\mathcal{Y} = \frac{1}{2}[\mathcal{X}, \mathcal{Y}]$.

It is perhaps worth to make a further comment on the relation between ad_X, left invariance and isometry. Since $\nabla_X Y = \mathrm{ad}_X Y - A_X Y$, rewriting (2.7) as

$$\underbrace{X\langle Y, Z\rangle}_{(*)} = \langle \mathrm{ad}_X Y, Z\rangle + \langle Y, \mathrm{ad}_X Z\rangle + \underbrace{\langle A_X Y, Z\rangle + \langle Y, A_X Z\rangle}_{(**)} \tag{4.11}$$

we have:

(i) if Y, Z are left invariant, then $(*) = 0$

(ii) if X is a Killing vector, $(**) = 0$

(iii) if Y, Z are left invariant and X is a Killing vector, then ad_X is skew-symmetric with respect to $\langle \cdot, \cdot\rangle$, i.e. the following equivalent quantities

$$\langle \mathrm{ad}_X Y, Z\rangle + \langle Y, \mathrm{ad}_X Z\rangle = 2\langle U(Y, Z), X\rangle = -\langle \mathbb{I}^{-1}\left(\mathrm{ad}^*_Y \mathbb{I} Z + \mathrm{ad}^*_Z \mathbb{I} Y\right), X\rangle \tag{4.12}$$

are all vanishing.
A Lie group admits a bi-invariant metric if and only if $\mathrm{ad}_{\mathcal{X}}$ is skew-symmetric with respect to $\langle \cdot, \cdot\rangle$ for all $\mathcal{X} \in \mathfrak{g}$ ([17], Lemma 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 4.4.

Given a vector field $\mathcal{X} \in \mathfrak{g}$, $\mathcal{X} = a^i A_i$ 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 $A_1, \ldots, A_n$: $x^i = \gamma e^{A_i}$ (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:

$$\begin{aligned}
\nabla_X Y &= (\nabla_X Y)^k B_k = \left(a^i \frac{\partial b^k}{\partial x^i} + a^i b^j \Gamma^k_{ij}\right) B_k \\
&= \left(\left(\mathcal{L}_X b^k\right) + a^i b^j \Gamma^k_{ij}\right)\gamma A_k = \gamma\left(\mathcal{L}_X \beta^k\right) A_k + \gamma a^i b^j \nabla_{A_i} A_j \qquad \text{by (4.3)} \\
&= \gamma\left(\left(\mathcal{L}_X b^k\right) A_k + \nabla_{\mathcal{X}}\mathcal{Y}\right)
\end{aligned} \tag{4.13}$$

Left invariance allows 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 $\frac{\partial}{\partial v^i} = \gamma A_i$, $i = 1, \ldots, n$. 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

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

(the first component living on $T_\gamma G$, the second on $T_{(\gamma,\gamma\mathcal{Y})}TG$). For matrix groups, $\nabla_{\dot\gamma}Y = 0$ corresponds to a linear frame being parallel transported along γ, see (2.15). Calling $\dot{\mathcal{X}} = \frac{\partial b^i}{\partial t}A_i$, in general, from (4.8), the covariant derivative is decomposed in the three parts:

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

Therefore a coordinate independent expression for X^h is

$$X^h = \left(\gamma\mathcal{X}\,;\; -\gamma(\frac{1}{2}[\mathcal{X},\, \mathcal{Y}] + U(\mathcal{X},\, \mathcal{Y}))\right) \tag{4.14}$$

whose integral curves are

$$\begin{aligned} \dot\gamma &= \gamma\mathcal{X} \\ \gamma\dot{\mathcal{Y}} &= -\gamma\nabla_{\mathcal{X}}\mathcal{Y} = -\gamma(\frac{1}{2}[\mathcal{X},\, \mathcal{Y}] + U(\mathcal{X},\, \mathcal{Y})) \end{aligned}$$

5. 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_gG, 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_gG and $T_eG = \mathfrak{g}$ that left (or right) translation implies. The three properties of a principal fiber bundle (see [11] 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 $\pi : TG \to \mathfrak{g}$ is the left translation itself and the fibers $\pi^{-1}(\mathcal{X})$, $\mathcal{X} \in \mathfrak{g}$, 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 $T_gG \to G \times \mathfrak{g}$, $(g,\, v_g) \mapsto (g,\, L_{g^{-1}\,*}v_g)$. 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. If the smooth manifold is a matrix group G, then the fiber bundle we are considering is normally referred to as G-structure (of G itself) and it is obtained by "reducing" $GL_n(\mathbb{R})$ to its subgroup G.

From (2.15), the condition $\nabla_{\dot\gamma}Y = 0$ allows 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 4.5 and 4.6, the compatibility condition between horizontal curves of $\mathbb{I}$ and horizontal curves of the fiber bundle structure of G (i.e. the "G-structure" of G itself) is that ∇ invariant to left translations and that $\nabla_{\mathcal{X}}\mathcal{X} = 0$. In fact, invariance plus $\nabla_{\mathcal{X}}\mathcal{X} = 0$ means that the parallel displacement can be carried out by left translations regardless of the path to follow.

The appearing of a nonvanishing term $U(\mathcal{X},\, \mathcal{Y})$ 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 (4.8), this happens whenever $\nabla_{\mathcal{X}}\mathcal{Y}$ is not completely skew symmetric.

So for the horizontal lift (4.14) the $T_{(\gamma,\gamma\mathcal{Y})}TG$ term splits into the horizontal component (in the fiber bundle) $-\gamma\frac{1}{2}[\mathcal{X},\, \mathcal{Y}]$ and the vertical one $-\gamma U(\mathcal{X},\, \mathcal{Y})$.

6. 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 [16]. 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 [6].

Left invariance allows 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:

$$\begin{aligned} T(0,\,t) = \dot\gamma(t) = \mathcal{G}_0(t)\mathcal{T}(t) = \gamma(t)\mathcal{T}(t) \\ S(0,\,t) = \delta\gamma(t) = \mathcal{G}_0(t)\mathcal{S}(t) = \gamma(t)\mathcal{S}(t) \end{aligned} \tag{6.1}$$

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 (2.10), 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 (4.13) becomes the following:

Proposition 6.1. *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:*

$$\begin{aligned} \nabla_{\dot\gamma}\delta\gamma = \gamma\left(\dot{\mathcal{S}} + \nabla_{\mathcal{T}}\mathcal{S}\right) \\ \nabla_{\delta\gamma}\dot\gamma = \gamma\left(\mathcal{T}' + \nabla_{\mathcal{S}}\mathcal{T}\right) \end{aligned} \tag{6.2}$$

Proof.

$$\begin{aligned} \nabla_{\dot\gamma}\delta\gamma &= \nabla_{\gamma\mathcal{T}(t)}\gamma\mathcal{S}(t) = \nabla_{\alpha^i(t)B_i}\beta^j(t)B_j \\ &= \left(\mathcal{L}_{\dot\gamma}\beta^j\right)B_j + \alpha^i\beta^j\nabla_{B_i}B_j = \frac{\partial\beta^j}{\partial t}B_j + \gamma\alpha^i\beta^j\nabla_{A_i}A_j \qquad \text{by (4.3)} \\ &= \gamma\frac{\partial\beta^j}{\partial t}A_j + \frac{1}{2}\gamma\alpha^i\beta^j\left([A_i,A_j] - \mathbb{I}^{-1}\mathrm{ad}^*_{A_i}\mathbb{I}A_j - \mathbb{I}^{-1}\mathrm{ad}^*_{A_j}\mathbb{I}A_i\right) \\ &= \gamma\frac{\partial\beta^j}{\partial t}A_j + \frac{1}{2}\gamma\left([\mathcal{T},\mathcal{S}] - \mathbb{I}^{-1}\mathrm{ad}^*_{\mathcal{T}}\mathbb{I}\mathcal{S} - \mathbb{I}^{-1}\mathrm{ad}^*_{\mathcal{S}}\mathbb{I}\mathcal{T}\right) = \gamma\left(\dot{\mathcal{S}} + \nabla_{\mathcal{T}}\mathcal{S}\right) \end{aligned}$$

and similarly

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

$\square$

The presence of two terms in both the covariant derivatives (6.2) 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 6.2. **(Symmetry lemma for reduction by group invariance in the Riemannian case)** *In the case of group manifold with Riemannian connection, the Lie algebra valued mixed derivatives are related by*

$$\mathcal{T}' = \dot{\mathcal{S}} \tag{6.3}$$

where $\mathcal{T}$ and $\mathcal{S}$ are respectively the Lie algebra valued infinitesimal variations for $\dot{\gamma}$ and $\delta\gamma$ defined in (6.1). *Furthermore*

$$\nabla_{\mathcal{T}}\mathcal{S} = \nabla_{\mathcal{S}}\mathcal{T} \tag{6.4}$$

Proof. From $\nabla_{\dot{\gamma}}\delta\gamma = \nabla_{\delta\gamma}\dot{\gamma}$ and the expression (6.2) for the two covariant derivatives:

$$\mathcal{T}' = \dot{\mathcal{S}} + [\mathcal{T},\, \mathcal{S}] \tag{6.5}$$

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 (6.4) follows. □

Notice that this is true only because we have chosen a torsion-free connection. In general when a non-Riemannian connection is chosen on the Lie group (for example the $(+)$ or $(-)$ canonical connections of Cartan), the "covariant infinitesimal variations" for the reduced principle have the more general expression (6.5), see [15] 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. The isometry (4.4) which correspond to parallel transport for ∇ along γ for left invariant vector fields splits after the reduction into two types of Lie algebra valued infinitesimal covariant variations, those indicated by " $\dot{}$ " (or " $'$ ") and those by the covariant derivative symbol. For left invariant vector fields $\mathcal{Y},\, \mathcal{Z} \in \mathfrak{g}$ along the curve γ of tangent vector $\dot{\gamma} = \gamma\mathcal{T}$, abusing notation one could write:

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

where $\langle \mathcal{Y},\dot{}\, \mathcal{Z}\rangle = \langle \dot{\mathcal{Y}},\, \mathcal{Z}\rangle + \langle \mathcal{Y},\, \dot{\mathcal{Z}}\rangle$ 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 (4.4).

PROPOSITION 6.3. *Given the left invariant vector fields $\mathcal{T},\, \mathcal{Y},\, \mathcal{Z} \in \mathfrak{g}$, the equation* (4.4) *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:*

$$\langle \dot{\mathcal{Y}},\, \mathcal{Z}\rangle = -\langle \mathcal{Y},\, \dot{\mathcal{Z}}\rangle \tag{6.6}$$

$$\langle \nabla_{\mathcal{T}}\mathcal{Y},\, \mathcal{Z}\rangle = -\langle \mathcal{Y},\, \nabla_{\mathcal{T}}\mathcal{Z}\rangle \tag{6.7}$$

Proof. Straightforward calculation from (2.11):

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

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 (4.4). ∎

COROLLARY 6.4. *For the family of variations $\mathcal{G}(s,\,t)$ of γ, with the notation above,*

$$\langle\dot{\mathcal{S}},\,\mathcal{T}\rangle = -\langle\mathcal{S},\,\dot{\mathcal{T}}\rangle \tag{6.8}$$

$$\langle\nabla_{\mathcal{T}}\mathcal{S},\,\mathcal{T}\rangle = -\langle\nabla_{\mathcal{T}}\mathcal{T},\,\mathcal{S}\rangle \tag{6.9}$$

$$\langle\nabla_{\mathcal{S}}\mathcal{S},\,\mathcal{T}\rangle = -\langle\mathcal{S},\,\nabla_{\mathcal{S}}\mathcal{T}\rangle \tag{6.10}$$

When the homogeneous space is naturally reductive, $\nabla_{\mathcal{T}}\mathcal{T}=0$ implies in (6.9) that also $\langle\nabla_{\mathcal{T}}\mathcal{S},\,\mathcal{T}\rangle=0$ and therefore the situation is much simpler.

REMARK 1. $\nabla_{\mathcal{T}}\mathcal{S}=\nabla_{\mathcal{S}}\mathcal{T}$ *is symmetric and belongs to V. In fact, from* (4.10) *and* $[\mathcal{T},\,\mathcal{S}]=0$

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

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

REMARK 2. *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 [16]) although it is normally not obtained using exclusively the tools from Riemannian geometry as we do here.

THEOREM 6.5. **(Reduced Hamilton principle)** *For $(G,\,\mathbb{I})$, the critical curves of the left invariant energy functional $\mathcal{E}=\int_{t_0}^{t_f}\langle\mathcal{T},\,\mathcal{T}\rangle dt$, 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*

$$\dot{\mathcal{T}}=-\nabla_{\mathcal{T}}\mathcal{T}=\mathbb{I}^{-1}\mathrm{ad}^*_{\mathcal{T}}\mathbb{I}\mathcal{T} \tag{6.11}$$

Proof. The proof can be obtained directly by inserting into the first variation formula (2.12) the value of the covariant derivative (6.2). Likewise, going through the reduction of the functional $\mathcal{E}$:

$$\begin{aligned}\frac{d}{ds}\mathcal{E}\left(\mathcal{G}_s(t)\right)\Big|_{s=0} &= \int_{t_0}^{t_f}\langle\nabla_S T,\,T\rangle dt\Big|_{s=0} = \int_{t_0}^{t_f}\langle\mathcal{T}'+\nabla_{\mathcal{S}}\mathcal{T},\mathcal{T}\rangle dt\\ &= -\int_{t_0}^{t_f}\langle\dot{\mathcal{T}},\,\mathcal{S}\rangle dt-\int_{t_0}^{t_f}\langle\nabla_{\mathcal{T}}\mathcal{T},\,\mathcal{S}\rangle dt \qquad \text{by (6.8) and (6.9)}\\ &= -\int_{t_0}^{t_f}\langle\dot{\mathcal{T}}+\nabla_{\mathcal{T}}\mathcal{T},\,\mathcal{S}\rangle dt\end{aligned}$$

□

The geodesic spray whose integral curves are the Euler-Poincaré equations (6.11) 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 (6.11) have only a left hand side: $\dot\gamma = \gamma\mathcal{T},\ \dot{\mathcal{T}} = 0$.

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 comes as natural simplification also for the control problems. For example, the forced second order vector field (2.18) reduces to

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

7. Reduction of the second order variational problem. When reducing higher order covariant derivatives, the two components of the covariant derivative mentioned in the previous Section mix up. For example, the "$\cdot$" part, applied to $\nabla_{\mathcal{X}}\mathcal{Y}$ results into:

$$(\dot{\nabla_{\mathcal{X}}\mathcal{Y}}) = \frac{\partial a^i b^j}{\partial t}\nabla_{A_i}A_j = \left(\frac{\partial a^i}{\partial t}b^j + a^i\frac{\partial b^j}{\partial t}\right)\nabla_{A_i}A_j = \nabla_{\dot{\mathcal{X}}}\mathcal{Y} + \nabla_{\mathcal{X}}\dot{\mathcal{Y}} \tag{7.1}$$

A few useful relations are:

Proposition 7.1. *For the family of variations $\mathcal{G}(s,\,t)$ on $(G,\,\mathbb{I})$:*

$$\nabla_{\delta\gamma}\nabla_{\dot\gamma}\dot\gamma = \gamma\left((\dot{\mathcal{T}})' + \nabla_{\mathcal{S}}\dot{\mathcal{T}} + \nabla_{\dot{\mathcal{S}}}\mathcal{T} + \nabla_{\mathcal{T}}\dot{\mathcal{S}} + \nabla_{\mathcal{S}}\nabla_{\mathcal{T}}\mathcal{T}\right) \tag{7.2}$$

$$\nabla_{\dot\gamma}\nabla_{\delta\gamma}\dot\gamma = \gamma\left((\dot{\mathcal{T}'}) + \nabla_{\mathcal{S}}\dot{\mathcal{T}} + \nabla_{\dot{\mathcal{S}}}\mathcal{T} + \nabla_{\mathcal{T}}\dot{\mathcal{S}} + \nabla_{\mathcal{T}}\nabla_{\mathcal{S}}\mathcal{T}\right) \tag{7.3}$$

$$\nabla^2{}_{\dot\gamma}\dot\gamma = \gamma\left(\ddot{\mathcal{T}} + 2\nabla_{\mathcal{T}}\dot{\mathcal{T}} + \nabla_{\dot{\mathcal{T}}}\mathcal{T} + \nabla^2{}_{\mathcal{T}}\mathcal{T}\right) \tag{7.4}$$

$$\nabla^3{}_{\dot\gamma}\dot\gamma = \gamma\Big(\dddot{\mathcal{T}} + 3\nabla_{\mathcal{T}}\ddot{\mathcal{T}} + 3\nabla_{\dot{\mathcal{T}}}\dot{\mathcal{T}} + \nabla_{\ddot{\mathcal{T}}}\mathcal{T} + 3\nabla^2{}_{\mathcal{T}}\dot{\mathcal{T}} + \tag{7.5}$$

$$+2\nabla_{\mathcal{T}}\nabla_{\dot{\mathcal{T}}}\mathcal{T} + \nabla_{\dot{\mathcal{T}}}\nabla_{\mathcal{T}}\mathcal{T} + \nabla^3{}_{\mathcal{T}}\mathcal{T}\Big)$$

$$R\,(\mathcal{S},\,\mathcal{T})\,\mathcal{T} = \nabla_{\mathcal{S}}\nabla_{\mathcal{T}}\mathcal{T} - \nabla_{\mathcal{T}}\nabla_{\mathcal{S}}\mathcal{T} = [\nabla_{\mathcal{S}},\,\nabla_{\mathcal{T}}]\mathcal{T} \tag{7.6}$$

Proof. We only prove (7.2), the other calculations being similar.

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

Concerning (7.6), from (2.9) with $[\delta\gamma,\, \dot\gamma] = [\mathcal{S},\, \mathcal{T}] = 0$, from (7.2) and (7.3)

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

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 $\mathfrak{g}$

$$(\dot{\mathcal{T}'}) = \frac{\partial^2\alpha^j}{\partial s\partial t}A_j = \frac{\partial^2\alpha^j}{\partial t\partial s}A_j = (\dot{\mathcal{T}})' \tag{7.7}$$

Hence, the result. □

The left invariance of the curvature tensor R means that the curvature term of (3.2) is:

$$R(\nabla_{\dot\gamma}\dot\gamma,\, \dot\gamma)\dot\gamma = R(\dot{\mathcal{T}} + \nabla_{\mathcal{T}}\mathcal{T},\, \mathcal{T})\gamma\mathcal{T} = \gamma\left(R(\dot{\mathcal{T}},\, \mathcal{T})\mathcal{T} + R(\nabla_{\mathcal{T}}\mathcal{T},\, \mathcal{T})\mathcal{T}\right) \tag{7.8}$$

The expressions (7.5) and (7.8) allows to write down directly the reduced expression for (3.2) 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 7.2. *For the family of variations $\mathcal{G}(s,\, t)$ on $(G,\, \mathbb{I})$:*

$$\langle(\dot{\mathcal{T}})',\, \dot{\mathcal{T}}\rangle = \langle\mathcal{S},\, \dddot{\mathcal{T}}\rangle \tag{7.9}$$
$$\langle(\dot{\mathcal{T}})',\, \nabla_{\mathcal{T}}\mathcal{T}\rangle = \langle\mathcal{S},\, \nabla_{\ddot{\mathcal{T}}}\mathcal{T} + 2\nabla_{\dot{\mathcal{T}}}\dot{\mathcal{T}} + \nabla_{\mathcal{T}}\ddot{\mathcal{T}}\rangle \tag{7.10}$$
$$\langle\nabla_{\mathcal{S}}\nabla_{\mathcal{T}}\mathcal{T},\, \dot{\mathcal{T}}\rangle = \langle\mathcal{S},\, \nabla^2_{\mathcal{T}}\dot{\mathcal{T}} + R(\dot{\mathcal{T}},\, \mathcal{T})\mathcal{T}\rangle \tag{7.11}$$
$$\langle\nabla_{\mathcal{S}}\nabla_{\mathcal{T}}\mathcal{T},\, \nabla_{\mathcal{T}}\mathcal{T}\rangle = \langle\mathcal{S},\, \nabla^3_{\mathcal{T}}\mathcal{T} + R(\nabla_{\mathcal{T}}\mathcal{T},\, \mathcal{T})\mathcal{T}\rangle \tag{7.12}$$
$$\langle\nabla_{\mathcal{T}}\dot{\mathcal{S}},\, \dot{\mathcal{T}}\rangle = \langle\mathcal{S},\, \nabla_{\dot{\mathcal{T}}}\dot{\mathcal{T}} + \nabla_{\mathcal{T}}\ddot{\mathcal{T}}\rangle \tag{7.13}$$
$$\langle\nabla_{\mathcal{T}}\dot{\mathcal{S}},\, \nabla_{\mathcal{T}}\mathcal{T}\rangle = \langle\mathcal{S},\, \nabla_{\dot{\mathcal{T}}}\nabla_{\mathcal{T}}\mathcal{T} + \nabla_{\mathcal{T}}\nabla_{\dot{\mathcal{T}}}\mathcal{T} + \nabla^2_{\mathcal{T}}\dot{\mathcal{T}}\rangle \tag{7.14}$$
$$\langle\nabla_{\mathcal{S}}\dot{\mathcal{T}},\, \dot{\mathcal{T}}\rangle = 0 \tag{7.15}$$
$$\langle\nabla_{\mathcal{S}}\dot{\mathcal{T}},\, \nabla_{\mathcal{T}}\mathcal{T}\rangle = -\langle\dot{\mathcal{T}},\, \nabla_{\mathcal{S}}\nabla_{\mathcal{T}}\mathcal{T}\rangle \tag{7.16}$$
$$\langle\nabla_{\dot{\mathcal{S}}}\mathcal{T},\, \dot{\mathcal{T}}\rangle = \langle\mathcal{S},\, \nabla_{\mathcal{T}}\ddot{\mathcal{T}}\rangle \tag{7.17}$$
$$\langle\nabla_{\dot{\mathcal{S}}}\mathcal{T},\, \nabla_{\mathcal{T}}\mathcal{T}\rangle = \langle\mathcal{S},\, \nabla_{\mathcal{T}}\nabla_{\dot{\mathcal{T}}}\mathcal{T} + 2\nabla^2_{\mathcal{T}}\dot{\mathcal{T}} + R(\dot{\mathcal{T}},\, \mathcal{T})\mathcal{T}\rangle \tag{7.18}$$

Proof. All the expressions are based on the equations (6.6) and (6.7). We see some of the significant calculations:

- Eq. (7.9):

$$\langle(\dot{\mathcal{T}})',\, \dot{\mathcal{T}}\rangle = \langle(\dot{\mathcal{T}'}),\, \dot{\mathcal{T}}\rangle = -\langle\mathcal{T}',\, \ddot{\mathcal{T}}\rangle = -\langle\dot{\mathcal{S}},\, \ddot{\mathcal{T}}\rangle = \langle\mathcal{S},\, \dddot{\mathcal{T}}\rangle \qquad \text{by (7.7)}$$

- Eq. (7.10):

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

- Eq. (7.11):

$$\begin{aligned}\langle \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{T}, \dot{\mathcal{T}} \rangle &= \langle \nabla_{\mathcal{T}} \nabla_{\mathcal{S}} \mathcal{T} + R(\mathcal{S}, \mathcal{T})\mathcal{T}, \dot{\mathcal{T}} \rangle \\ &= -\langle \nabla_{\mathcal{S}} \mathcal{T}, \nabla_{\mathcal{T}} \dot{\mathcal{T}} \rangle + \langle R(\mathcal{S}, \mathcal{T})\mathcal{T}, \dot{\mathcal{T}} \rangle \qquad \text{by (7.6)} \\ &= -\langle \nabla_{\mathcal{T}} \mathcal{S}, \nabla_{\mathcal{T}} \dot{\mathcal{T}} \rangle + \langle R(\dot{\mathcal{T}}, \mathcal{T})\mathcal{T}, \mathcal{S} \rangle \qquad \text{by (6.4) and (3.3)} \\ &= \langle \mathcal{S}, \nabla^2_{\mathcal{T}} \dot{\mathcal{T}} \rangle + \langle R(\dot{\mathcal{T}}, \mathcal{T})\mathcal{T}, \mathcal{S} \rangle\end{aligned}$$

- Eq. (7.17):

$$\begin{aligned}\langle \nabla_{\dot{\mathcal{S}}} \mathcal{T}, \dot{\mathcal{T}} \rangle &= \langle (\dot{\nabla_{\mathcal{S}} \mathcal{T}}) - \nabla_{\mathcal{S}} \dot{\mathcal{T}}, \dot{\mathcal{T}} \rangle = -\langle \nabla_{\mathcal{S}} \mathcal{T}, \ddot{\mathcal{T}} \rangle - \langle \nabla_{\mathcal{S}} \dot{\mathcal{T}}, \dot{\mathcal{T}} \rangle \qquad \text{by (7.1)} \\ &= -\langle \nabla_{\mathcal{T}} \mathcal{S}, \ddot{\mathcal{T}} \rangle = \langle \mathcal{S}, \nabla_{\mathcal{T}} \ddot{\mathcal{T}} \rangle \qquad \text{by (6.4) and (7.15)}\end{aligned}$$

The other relations are obtained using similar arguments. □

THEOREM 7.3. *A necessary condition for a smooth curve $\gamma(t) \in G$ of tangent vector field $\mathcal{T} \in \mathfrak{g}$ 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*

$$\dddot{\mathcal{T}} + 3\nabla_{\mathcal{T}} \ddot{\mathcal{T}} + 3\nabla_{\dot{\mathcal{T}}} \dot{\mathcal{T}} + \nabla_{\ddot{\mathcal{T}}} \mathcal{T} + 3\nabla^2{}_{\mathcal{T}} \dot{\mathcal{T}} + 2\nabla_{\mathcal{T}} \nabla_{\dot{\mathcal{T}}} \mathcal{T} + \nabla_{\dot{\mathcal{T}}} \nabla_{\mathcal{T}} \mathcal{T} + \nabla^3{}_{\mathcal{T}} \mathcal{T} + R(\dot{\mathcal{T}}, \mathcal{T})\mathcal{T} + R(\nabla_{\mathcal{T}} \mathcal{T}, \mathcal{T})\mathcal{T} = 0 \tag{7.19}$$

Proof. From (7.5) and (7.8) we have already (7.19) by brute force. To see it directly, expand the expression for $\mathcal{J}$ and use the computations of Proposition 7.2.

$$\begin{aligned}\frac{d}{ds} \mathcal{J}\left(\mathcal{G}_s(t)\right)\Big|_{s=0} &= \int_{t_0}^{t_f} \langle \nabla_{\delta\gamma} \nabla_{\dot{\gamma}} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle dt \\ = \int_{t_0}^{t_f} \langle \gamma \left((\dot{\mathcal{T}})' + \nabla_{\mathcal{S}} \dot{\mathcal{T}} + \nabla_{\dot{\mathcal{S}}} \mathcal{T} + \nabla_{\mathcal{T}} \dot{\mathcal{S}} + \nabla_{\mathcal{S}} \nabla_{\mathcal{T}} \mathcal{T} \right), \gamma \left(\dot{\mathcal{T}} + \nabla_{\mathcal{T}} \mathcal{T} \right) \rangle dt \qquad &\text{by (7.2)}\end{aligned}$$

The ten inner products under the sign of integral are given in (7.9)-(7.18). The result follows by considering extremals of $\mathcal{J}$ i.e. $\frac{d}{ds} \mathcal{J}\left(\mathcal{G}_s(t)\right)\big|_{s=0} = 0$. □

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.

8. Optimal control for the reduced second order variational problem. Assume that the left invariant input distribution $\mathcal{F}$ has the coordinate expression $\mathcal{F} = f^i A_i$.

PROPOSITION 8.1. *The covariant derivatives of the input vector distribution $\mathcal{F}$ are:*

$$\nabla_{\dot{\gamma}} \gamma \mathcal{F} = \nabla^2{}_{\dot{\gamma}} \dot{\gamma} = \gamma \left(\dot{\mathcal{F}} + \nabla_{\mathcal{T}} \mathcal{F} \right) \tag{8.1}$$

$$\nabla^2_{\dot{\gamma}} \gamma \mathcal{F} = \nabla^3{}_{\dot{\gamma}} \dot{\gamma} = \gamma \left(\ddot{\mathcal{F}} + 2\nabla_{\mathcal{T}} \dot{\mathcal{F}} + \nabla_{\dot{\mathcal{T}}} \mathcal{F} + \nabla^2{}_{\mathcal{T}} \mathcal{F} \right) \tag{8.2}$$

Proof. Equation (8.2) is shown in the same way as (6.2). For (8.2), it can be obtained from computations of the same type as in Proposition 7.1.

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

□

Therefore, the expression for the control $\mathcal{F}$ is obtained from (3.4). Resuming the "C^2 dynamic interpolation problem" of Section 3:

PROPOSITION 8.2. *Under the same assumptions of Proposition 3.2, 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*

$$\ddot{\mathcal{F}} + 2\nabla_{\mathcal{T}}\dot{\mathcal{F}} + \nabla_{\dot{\mathcal{T}}}\mathcal{F} + \nabla^2{}_{\mathcal{T}}\mathcal{F} + R(\mathcal{F},\,\mathcal{T})\mathcal{T} = 0 \qquad \textit{subject to} \qquad \dot{\mathcal{T}} + \nabla_{\mathcal{T}}\mathcal{T} = \mathcal{F} \tag{8.3}$$

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

Proof. Since R is a left invariant tensor, this is the straightforward substitution in (3.4) of equation (8.2). □

9. Application to $SE(3)$. The *Special Euclidean Group $SE(3)$* is the Lie group of isometric transformations of $\mathbb{R}^3$ i.e. its left (or right) action on $\mathbb{R}^3$

$$
\begin{aligned}
SE(3)\times\mathbb{R}^3 &\to \mathbb{R}^3\\
(g,\,x) &\mapsto L_g x = gx
\end{aligned}
$$

has push forward

$$
\begin{aligned}
SE(3)\times T_x\mathbb{R}^3 = \mathbb{R}^3 &\to T_{gx}\mathbb{R}^3 = \mathbb{R}^3\\
(g,\,v) &\mapsto L_{g*}v = gv
\end{aligned}
$$

which is an isometry and takes straight lines to straight lines in $\mathbb{R}^3$.

Using homogeneous coordinates,

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

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

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

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

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:

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

while in $SE(3)$:

$$\begin{array}{rccc} \mathrm{e}\ : & \mathfrak{se}(3) & \to & SE(3) \\ & \mathcal{X} = \left[\begin{array}{cc} \widehat{\omega}_{\mathcal{X}} & v_{\mathcal{X}} \\ 0_{3\times 1} & 0 \end{array}\right] & \mapsto & \left[\begin{array}{cc} \mathrm{e}^{\widehat{\omega}_{\mathcal{X}}} & A(\widehat{\omega}_{\mathcal{X}})v_{\mathcal{X}} \\ 0 & 1 \end{array}\right] \end{array} \tag{9.1}$$

where

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

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

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

gives the *Lie bracket* $\mathrm{ad}_{\mathcal{X}}(\mathcal{Y}) = [\mathcal{X}, \mathcal{Y}] = \mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X}$, i.e. the bilinear form defining the Lie algebra. The Lie brackets basis elements $A_1, \ldots, A_6$ of $\mathfrak{se}(3)$ are expressed in terms of the *structural constants* c_{ij}^k : $[A_i, A_j] = c_{ij}^k A_k$. The linear representations of the operators $\mathrm{Ad}_g(\cdot)$ and $\mathrm{ad}_{\mathcal{X}}(\cdot)$ is:

$$\mathrm{Ad}_g = \begin{bmatrix} R & 0 \\ \widehat{p}R & R \end{bmatrix} \qquad \mathrm{ad}_{\mathcal{X}} = \begin{bmatrix} \widehat{\omega}_{\mathcal{X}} & 0 \\ \widehat{v}_{\mathcal{X}} & \widehat{\omega}_{\mathcal{X}} \end{bmatrix} \tag{9.2}$$

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 [5]. 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 [5, 10], 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 [19] §13).

For the metric structure we are adopting, the Riemannian exponential map Exp differs from the Lie group exponential map (9.1). In fact, disregarding the action $SO(3) \to \mathrm{End}(\mathbb{R}^3)$ means dropping the $A(\widehat{\omega})$ term:

$$\begin{array}{rccc} \mathrm{Exp}\ : & \mathfrak{se}(3) & \to & SE(3) \\ & \mathcal{X} = \left[\begin{array}{cc} \widehat{\omega}_{\mathcal{X}} & v_{\mathcal{X}} \\ 0_{3\times 1} & 0 \end{array}\right] & \mapsto & \left[\begin{array}{cc} \mathrm{e}^{\widehat{\omega}_{\mathcal{X}}} & v_{\mathcal{X}} \\ 0 & 1 \end{array}\right] \end{array} \tag{9.3}$$

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 (9.2), the expressions for the coadjoint and infinitesimal coadjoint actions $\mathrm{Ad}^*_{g^{-1}}$ and $\mathrm{ad}^*_{\mathcal{X}}$ are:

$$\begin{aligned}\mathrm{Ad}^*_{g^{-1}} &= (\mathrm{Ad}_g)^{-T} = \begin{bmatrix} R & \widehat{p}R \\ 0 & R\end{bmatrix},\\ \mathrm{ad}^*_{\mathcal{X}} &= -\left.\frac{d}{dt}\mathrm{Ad}^*_{\mathrm{e}^{-t\mathcal{X}}}\right|_{t=0} = -\left(\mathrm{ad}_{\mathcal{X}}\right)^T = \begin{bmatrix} -\widehat{\omega}_{\mathcal{X}} & -\widehat{v}_{\mathcal{X}} \\ 0 & -\widehat{\omega}_{\mathcal{X}}\end{bmatrix}\end{aligned}$$

In (4.10), 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 $\mathrm{ad}^*_{\mathcal{X}}\mathbb{I}\mathcal{Y}$ are nonnull, even when the inertia tensor is diagonal, $\mathbb{I}=I$. In this case $\mathbb{I}$ can be pulled out and (with abuse of notation)

$$\mathrm{ad}^*_{\mathcal{X}}\mathbb{I}\mathcal{Y} = \mathrm{ad}^*_{\mathcal{X}}\mathcal{Y} = \begin{bmatrix} 0 \\ -\widehat{\omega}_{\mathcal{X}} v_{\mathcal{Y}}\end{bmatrix} = \begin{bmatrix} 0 \\ v_{\mathcal{Y}}\times\omega_{\mathcal{X}}\end{bmatrix} \neq 0$$

In particular then

$$U(\mathcal{X},\,\mathcal{Y}) = -\frac{1}{2}\left(\mathrm{ad}^*_{\mathcal{X}}\mathcal{Y} + \mathrm{ad}^*_{\mathcal{Y}}\mathcal{X}\right) = \frac{1}{2}\begin{bmatrix} 0 \\ \widehat{\omega}_{\mathcal{X}} v_{\mathcal{Y}} + \widehat{\omega}_{\mathcal{Y}} v_{\mathcal{X}}\end{bmatrix} \in \mathbb{R}^3 \tag{9.4}$$

Since $\mathrm{ad}_{\mathcal{X}}\mathcal{Y} = \begin{bmatrix} \widehat{\omega}_{\mathcal{X}}\omega_{\mathcal{Y}} \\ \widehat{\omega}_{\mathcal{X}} v_{\mathcal{Y}} - \widehat{\omega}_{\mathcal{Y}} v_{\mathcal{X}}\end{bmatrix}$, the covariant derivative is

$$\nabla_{\mathcal{X}}\mathcal{Y} = \frac{1}{2}\mathrm{ad}_{\mathcal{X}}\mathcal{Y} + U(\mathcal{X},\,\mathcal{Y}) = \begin{bmatrix} \frac{1}{2}\widehat{\omega}_{\mathcal{X}}\omega_{\mathcal{Y}} \\ \widehat{\omega}_{\mathcal{X}} v_{\mathcal{Y}}\end{bmatrix} \tag{9.5}$$

The linear map $A_{\mathcal{Y}}$ of Proposition 4.1 is

$$A_{\mathcal{Y}} = \begin{bmatrix} \frac{1}{2}\widehat{\omega}_{\mathcal{Y}} & 0 \\ \widehat{v}_{\mathcal{Y}} & 0\end{bmatrix}$$

which is skew-symmetric with respect to $\langle\,\cdot\,,\,\cdot\,\rangle$ for all $\mathcal{X}$ if and only if $\mathcal{Y}\in\mathfrak{so}(3)$. Therefore $\mathcal{Y}\not\in\mathfrak{so}(3)$ 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 $\mathcal{Y}=\begin{bmatrix}\omega_{\mathcal{Y}} \\ 0\end{bmatrix}$. In fact, $\mathcal{Y}\in\mathfrak{so}(3)$ implies that $\mathrm{e}^{\mathcal{Y}}=\mathrm{Exp}(\mathcal{Y})$ and therefore one parameter subgroups generated by $\mathcal{Y}$ coincide with geodesics of $\mathbb{I}$. Notice from (9.4) 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:

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

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 (6.11))

$$\dot{\mathcal{X}} = \begin{bmatrix}\dot{\omega}_{\mathcal{X}} \\ \dot{v}_{\mathcal{X}}\end{bmatrix} = -\nabla_{\mathcal{X}}\mathcal{X} = \begin{bmatrix} 0 \\ -\widehat{\omega}_{\mathcal{X}} v_{\mathcal{X}}\end{bmatrix}$$

with geodesic spray

$$\Gamma = \gamma\mathcal{X}^h = \left(\gamma\mathcal{X}\,;\ -\gamma U(\mathcal{X},\,\mathcal{X})\right) = \left(\gamma\begin{bmatrix}\omega_{\mathcal{X}} \\ v_{\mathcal{X}}\end{bmatrix}\ ;\ \ \gamma\begin{bmatrix} 0 \\ \widehat{\omega}_{\mathcal{X}} v_{\mathcal{X}}\end{bmatrix}\right)$$

Similarly to (9.5), we have

$$\nabla_{\dot{\mathcal{X}}}\mathcal{Y} = \begin{bmatrix} \frac{1}{2}\dot{\widehat{\omega}}_{\mathcal{X}}\omega_{\mathcal{Y}} \\ \dot{\widehat{\omega}}_{\mathcal{X}}v_{\mathcal{Y}} \end{bmatrix}, \qquad \nabla_{\mathcal{X}}\dot{\mathcal{Y}} = \begin{bmatrix} \frac{1}{2}\widehat{\omega}_{\mathcal{X}}\dot{\omega}_{\mathcal{Y}} \\ \widehat{\omega}_{\mathcal{X}}\dot{v}_{\mathcal{Y}} \end{bmatrix}, \qquad \nabla_{\mathcal{X}}\nabla_{\mathcal{Y}}\mathcal{Z} = \frac{1}{4}\begin{bmatrix} \widehat{\omega}_{\mathcal{X}}\widehat{\omega}_{\mathcal{Y}}\omega_{\mathcal{Z}} \\ 4\widehat{\omega}_{\mathcal{X}}\widehat{\omega}_{\mathcal{Y}}v_{\mathcal{Z}} \end{bmatrix} \tag{9.6}$$

and

$$\nabla_{\mathcal{W}}\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}}\mathcal{Z} = \frac{1}{8}\begin{bmatrix} \widehat{\omega}_{\mathcal{W}}\widehat{\omega}_{\mathcal{X}}\widehat{\omega}_{\mathcal{Y}}\omega_{\mathcal{Z}} \\ 8\widehat{\omega}_{\mathcal{W}}\widehat{\omega}_{\mathcal{X}}\widehat{\omega}_{\mathcal{Y}}v_{\mathcal{Z}} \end{bmatrix}$$

From (2.9), 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 by[1]

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

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

$$\begin{aligned} r_2 = {} & \frac{3}{4}\omega_{\mathcal{X}} \times \omega_{\mathcal{Y}} \times v_{\mathcal{Z}} + \frac{3}{4}\omega_{\mathcal{Y}} \times \omega_{\mathcal{X}} \times v_{\mathcal{Z}} + \frac{1}{4}(\omega_{\mathcal{X}} \times \omega_{\mathcal{Y}}) \times v_{\mathcal{Z}} \\ & + \frac{1}{4}\omega_{\mathcal{X}} \times \omega_{\mathcal{Z}} \times v_{\mathcal{Y}} + \frac{3}{4}\omega_{\mathcal{Z}} \times \omega_{\mathcal{X}} \times v_{\mathcal{Y}} - \frac{1}{4}(\omega_{\mathcal{X}} \times \omega_{\mathcal{Z}}) \times v_{\mathcal{Y}} \\ & + \frac{1}{4}\omega_{\mathcal{Y}} \times \omega_{\mathcal{Z}} \times v_{\mathcal{X}} - \frac{3}{4}\omega_{\mathcal{Z}} \times \omega_{\mathcal{Y}} \times v_{\mathcal{X}} - \frac{1}{4}(\omega_{\mathcal{Y}} \times \omega_{\mathcal{Z}}) \times v_{\mathcal{X}} \end{aligned}$$

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

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

- if $\mathcal{X},\mathcal{Y} \in \mathbb{R}^3$ $\qquad R = 0$
- 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, 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 [1, 12], the general expression for the (nonnormalized) two-plane curvature for a semidirect product is given by:

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

therefore

[1] Warning: in [20] a wrong expression is reported

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

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

- if $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^3 \quad R \equiv 0$
- 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

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

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 (7.19) corresponds to the system:

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

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 (2.17), eq. (8.3) becomes the system of ordinary differential equations which are linear and second order in $\mathcal{F}$, quadratic and first order in $\mathcal{T}$:

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

REFERENCES

[1] V. I. Arnold, *Mathematical methods of Classical Mechanics*, vol. 60 of Graduate texts in mathematics, Springer-Verlag, 2nd ed., 1989.

[2] F. Bullo, N. Leonard, and A. Lewis, *Controllability and motion algorithms for underactuates Lagrangian systems on Lie groups*, IEEE Trans. on Automatic Control, 45 (2000), pp. 1437–1454.

[3] F. Bullo and R. Murray, *Tracking for fully actuated mechanical systems: a geometric framework*, Automatica, 35 (1999), pp. 17–34.

[4] 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, pp. 1636–1641.

[5] E. Cartan, *La géometrie des groupes de transformations*, in Œuvres Complètes, vol. 2, part I, Gauthier-Villars, Paris, France, 1953, pp. 673–792.

[6] H. Cendra, D. Holm, J. Marsden, and T. Ratiu, *Lagrangian reduction, the Euler-Poincaré equations and semidirect products*, Amer. Math. Soc. Transl., 186 (1998), pp. 1–25.

[7] M. Crampin and F. Pirani, *Applicable differential geometry*, London Mathematical Society Lecture notes, Cambridge University Press, Cambridge, UK, 1986.

[8] P. E. Crouch and F. Silva Leite, *The dynamic interpolation problem on Riemannian manifolds, Lie groups and symmetric spaces*, Journal of Dynamical and Control Systems, 1 (1995), pp. 177–202.

[9] M. do Carmo, *Riemannian geometry*, Birkhäuser, Boston, 1992.

[10] L. Eisenhart, *Riemannian geometry*, Princeton University Press, Princeton, 1966.

[11] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, vol. I and vol. II, Interscience Publisher, New York, 1963 and 1969.

[12] B. B. L., *Sur la curbure des metriques Riemanniennes invariant des groups de lie et des espaces homogenes*, Ann. Sci. Ecole National Superior, 11 (1978), p. 543.

[13] A. Lewis and R. Murray, *Configuration controllability of simple mechanical control systems*, SIAM Journal on Control and Optimization, 35 (1997), pp. 766–790.

[14] ———, *Decompositions for control systems on manifolds with an affine connection*, Systems and Control Letters, 31 (1997), pp. 199–205.

[15] J. Marsden, *Lectures on Mechanics*, Cambridge University Press, Cambridge, 1992.

[16] J. Marsden and T. Ratiu, *Introduction to mechanics and symmetry*, vol. 17 of Texts in Applied Mathematics, Springer-Verlag, 2nd ed., 1999.

[17] J. Milnor, *Curvature of left invariant metrics on lie groups*, Advances in Mathematics, 21 (1976), pp. 293–329.

[18] L. Noakes, G. Heinzinger, and B. Paden, *Cubic splines on curved spaces*, IMA Journal of Mathematics, Control and Information, 12 (1989), pp. 465–473.

[19] K. Nomizu, *Invariant affine connections on homogeneous spaces*, American Journal of Mathematics, 76 (1954), pp. 33–65.

[20] M. Zefran, V. Kumar, and C. Croke, *On the generation of smooth three-dimensional rigid body motions*, IEEE Trans. on Robotics and Automation, 12 (1998), pp. 576–589.