Tutorial

Manifold Calculus in System Theory and Control—Second Order Structures and Systems

Simone Fiori

Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Via Brecce Bianche, I-60131 Ancona, Italy; s.fiori@staff.univpm.it

Abstract: The present tutorial paper constitutes the second of a series of tutorials on manifold calculus with applications in system theory and control. The aim of the present tutorial, in particular, is to explain and illustrate some key concepts in manifold calculus such as covariant derivation and manifold curvature. Such key concepts are then applied to the formulation, to the control, and to the analysis of non-linear dynamical systems whose state-space are smooth (Riemannian) manifolds. The main flow of exposition is enriched by a number of examples whose aim is to clarify the notation used and the main theoretical findings through practical calculations.

Keywords: second-order dynamical systems on manifold; feedback control system; smooth manifolds; covariant derivation; manifold curvature

1. Introduction

The present paper appears as the second part of a long tutorial covering notions of calculus on manifolds embedded in larger flat spaces, dynamical systems whose state spaces are structured as smooth manifolds and non-linear control systems on manifolds.

Regulation of dynamical systems on state manifolds appears to be an appealing research field in non-linear control theory which is gaining increasing interest in the scientific community, especially in the field of mechanical systems control [1–6]. Depending on the application at hand, one may be facing two cases [7]: positional control, arising in the regulation of first-order as well as second-order systems, and velocity control, arising in the regulation of second-order systems on manifolds. Embedded manifold setting proves useful in modeling large-scale constrained systems, since a state manifold allows one to take into account non-linear constraints on a non-linear system's variables seamlessly. A number of scientific papers, both of theoretic and applied nature, were published over the years to investigate such extensions. Since the early days of control theory, it was clear how geometric methods are of prime importance in the study of dynamical systems and their regulation [8–11]. For a short review of contributions in the theory of Lie-group as well as general manifold system regulation, interested readers might like to consult [7].

A challenging problem in feedback control on smooth manifolds is the synchronization in time of non-linear dynamical systems, an appealing application of which is in the synchronization of a follower oscillator, whose state-variables take the role of masking carrier, to a leader oscillator, to be employed in data encryption for secure transmission over digital networks [12]. In such application, it is assumed that the leader oscillator state variable is completely observable, that its output coincides with its state and that its dynamics is completely specified.

The present tutorial paper is devoted to describing and illustrating key concepts in manifold calculus such as covariant derivation and manifold curvature with applications to formulating, controlling and studying the properties of dynamical systems whose state-space is a smooth manifold. Its content may be summarized as follows:

- it provides a clear and well-motivated introduction to advanced concepts in manifold calculus, the basis of system and control theory on manifolds, with special emphasis to its computational and application aspects. The present contribution provides practical formulas to cope with those *real-valued* manifolds which, in the author's experience, are most accessed in engineering and applied science problems. As a matter of fact, *complex-valued* manifolds are not treated at all;
- it clearly states and illustrates the idea that, when dealing with a computer-based implementation of dynamical systems, it is necessary to discretize the differential equations that describe the system in a suitable way. In order to perform such discretization, standard discretization methods (such as the ones based on Euler forward-backward discretization) are not advisable since these do not work as they stand on curved manifolds. It is therefore advisable to resort to more sophisticated integration techniques such as the one based on geodesics.

The present paper is organized as follows. Section 2 is dedicated to a recapitulation of the notation used in the first part [13] as well as the present part, along with the main concepts covered by the first part. The recapitulation offered in Section 2 is not very detailed as readers that are approaching the present part of the tutorial are supposed to either get themselves sufficiently acquainted with the introductory material presented in the first part or to be already familiar with it. Section 3 introduces the key notion of covariant derivation of a vector field, which generalizes to manifold the familiar notion of directional derivative from real calculus. Section 3 discusses such concept in detail, starting from motivations behind its definition. Section 4 takes off by some observations about iterated covariant derivation and lands at the key concept of manifold curvature, which is a quintessential notion in manifold calculus. Section 5 surveys the notion of calculus of variations on manifolds by the help of a family of smooth curves and shows how to derive a general kind of continuous-time dynamical systems described by second-order differential equations on manifold. Control systems to regulate such second-order dynamical systems are discussed in Section 6 with special emphasis to time-synchronization. Section 7 presents an overview of second-order discrete-time systems on manifolds with special emphasis on ARMA-type random path generators. Section 8 concludes this tutorial.

As a distinctive aspect of the present tutorial, the main flow of discussion is based on *coordinate-free* (or component-free) expressions, which facilitates the implementation of the main equations by a matrix-friendly computation language such as MATLAB$^{©}$. Starred subsections, namely subsections marked by an asterisk (*), treat some specific arguments, some related to *coordinate-prone* manifold calculus and some of minor importance in the present context, which may be skipped by the uninterested readers without detriment to the comprehension of the main flow.

2. Notation and Recapitulation of Fundamentals

In the present section we recall some notation and the principal notions laid out in the first part [13]. For details on derivations, we point the reader to such introductory part.

In the present paper, a column array is denoted by a lower-case symbol, while a matrix is denoted by an upper-case symbol. An element of a manifold is however always denoted by a lower-case symbol.

It is convenient to recall a few manifolds of interest in application, defined in embedded (coordinate-free) terms:

- **Hypercube**: The simplest manifold of interest is the hypercube $\mathbb{R}^p$, which is essentially the set spanned by p real-valued variables (or p-tuples). A generalization of the hypercube is the set $\mathbb{R}^{p \times q}$ of real-valued $p \times q$ matrices.
- **Hypersphere**: A hypersphere is represented as $\mathbb{S}^{p-1} := \{x \in \mathbb{R}^p \mid x^\top x = 1\}$ and is the subset of points of the hypercube with unit Euclidean distance from the point 0. This is a smooth manifold of dimension $p - 1$.
- **General linear group and special linear group**: The general linear group is defined as $\mathbb{GL}(p) := \{X \in \mathbb{R}^{p \times p} \mid \det(X) \neq 0\}$. The special linear group is defined as

$Sl(p) := \{X \in \mathbb{R}^{p \times p} \mid \det(X) = 1\}$. It represents the set of invertible linear operators of given dimension.

- **Special orthogonal group**: The manifold of special orthogonal matrices is defined as $\mathbb{SO}(p) := \{X \in \mathbb{R}^{p \times p} \mid X^\top X = I_p, \det(X) = 1\}$. It represents the set of hyper-rotations in a hypercube of given dimension. In the definition, the symbol I_n denotes a $n \times n$ identity matrix.

- **Stiefel manifold**: The (compact) Stiefel manifold is defined as:

$$St(n, p) := \{X \in \mathbb{R}^{n \times p} \mid X^\top X = I_p\}, \tag{1}$$

where $p \leqslant n$.

- **Real symplectic group**: The real symplectic group is defined as

$$Sp(2n) := \{Q \in \mathbb{R}^{2n \times 2n} \mid Q^\top J Q = J\}, \quad J := \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}, \tag{2}$$

where the symbol 0_n denotes a whole-zero $n \times n$ matrix.

- **Manifold of symmetric, positive-definite (SPD) matrices**: The space of symmetric, positive-definite matrices is defined as

$$\mathbb{S}^+(p) := \{P \in \mathbb{R}^{p \times p} \mid P = P^\top, P > 0\}, \tag{3}$$

where the notation $P > 0$ indicates a positive-definite matrix.

- **Grassmann manifold**: A Grassmann manifold $Gr(n, p)$ is a set of subspaces of $\mathbb{R}^n$ spanned by p independent vectors, namely

$$Gr(n, p) = \{\operatorname{span}(w_1, w_2, w_3, \ldots, w_p)\}, \tag{4}$$

with $w_1, w_2, w_3, \ldots, w_p \in \mathbb{R}^n$ being a p-tuple of arbitrary and linearly independent n-dimensional arrays.

A fundamental assumption of the present tutorial is that any manifold $\mathbb{M}$ is embedded into an ambient space $\mathbb{A}$ of suitable dimension. The ambient space is assumed to be metric, namely to be associated with a scalar product $\langle \star, \star \rangle^\mathbb{A}$.

A smooth curve is generally denoted as $\gamma : [-\epsilon, \epsilon] \to \mathbb{M}$, with $\epsilon > 0$. A curve may be thought of as a trajectory generated by a dynamical system whose state-space is a smooth manifold.

The tangent space to the manifold $\mathbb{M}$ at a point x is denoted as $T_x\mathbb{M}$. The tangent bundle is a collection of tangent spaces, namely $T\mathbb{M} := \{(x, v) \in \mathbb{A}^2 \mid x \in \mathbb{M}, v \in T_x\mathbb{M}\}$. The concept of normal space complements that of tangent space, namely, the normal space of an embedded manifold in a given point under a chosen metric is defined as $N_x\mathbb{M} := \{z \in \mathbb{A} \mid \langle z, v \rangle^\mathbb{A} = 0, \forall v \in T_x\mathbb{M}\}$.

A vector field on a manifold $\mathbb{M}$ is a function $f : \mathbb{M} \to T\mathbb{M}$ that assigns a tangent vector $f(x) \in T_x\mathbb{M}$ to every point $x \in \mathbb{M}$. A vector field may even depend on the time parameter, in which case it will be denoted as $f(t, x)$ with $f : \mathbb{R} \times \mathbb{M} \to T\mathbb{M}$.

In order to take into account non-linear constraints on the state variables of a dynamical system, it is convenient to introduce the notion of *curved* state space under the form of smooth manifolds. This logical process leads to a type of first-order (or single-integrator) dynamical systems described by

$$\dot{x}(t) = f(t, x(t)), \quad x(0) = x_0 \in \mathbb{M}, \quad t \in [0, t_f], \tag{5}$$

where $f : \mathbb{R} \times \mathbb{M} \to T\mathbb{M}$ denotes the state-transition function of the system and $x \in \mathbb{M}$ denotes the state of the system at any given time t.

Given two Riemannian manifolds $\mathbb{M}$ and $\mathbb{N}$, let $x \in \mathbb{M}$: A pushforward map $\mathrm{d}_x f : T_x\mathbb{M} \to T_{f(x)}\mathbb{N}$ is defined such that for every smooth curve $\gamma_x : [-\epsilon, \epsilon] \to \mathbb{M}$, such that $\gamma_x(0) = x$, it holds:

$$\mathrm{d}_x f(\dot{\gamma}_x(0)) := \left.\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_x(t))\right|_{t=0}. \tag{6}$$

In general, given a function $f : \mathbb{M} \to \mathbb{N}$ that maps a point x from the manifold $\mathbb{M}$ to the point $f(x)$ on the manifold $\mathbb{N}$, the map $\mathrm{d}_x f$ associates to any tangent vector v belonging to the tangent space $T_x\mathbb{M}$ a tangent vector $\mathrm{d}_x f(v)$ belonging to the tangent space $T_{f(x)}\mathbb{N}$ by 'pushing' the former to the latter space.

Manifolds that also enjoy the distinguishing properties of an algebraic group are termed Lie groups. The tangent space of a Lie group $\mathbb{G}$ at the identity is termed Lie algebra and is denoted as $\mathfrak{g}$. Any Lie group admits an exponential map, denoted as $\exp : \mathfrak{g} \to \mathbb{G}$.

A Riemannian manifold is defined as a smooth manifold $\mathbb{M}$ whose tangent bundle is equipped with a smooth family of inner products $\mathbb{M} \ni x \mapsto \langle \star, \star \rangle_x \in \mathbb{R}$ termed 'metric'. There exists a canonical metric related to the ambient space $\mathbb{A}$ that a manifold $\mathbb{M}$ is embedded in, given by $\langle u, w \rangle_x = \langle u, G_x(w) \rangle^{\mathbb{A}}$, where $\langle \star, \star \rangle^{\mathbb{A}}$ denotes an inner product in $\mathbb{A}$. Such expression is based on a metric kernel $G : T\mathbb{M} \to T\mathbb{M}$. A derivative of the metric kernel with respect to a curve is defined by

$$G_\gamma^\bullet(w, \dot{\gamma}) := \frac{\mathrm{d}}{\mathrm{d}t} G_\gamma(w), \tag{7}$$

where γ denotes any smooth curve.

A geodesic arc is defined as the curve γ satisfying the following differential inclusion

$$\frac{\partial \mathscr{F}}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathscr{F}}{\partial \dot{\gamma}} \in N_\gamma\mathbb{M}. \tag{8}$$

where $\mathscr{F}(x, v) := \langle v, v \rangle_x$ denotes a fundamental form. A solution of such differential inclusion that joins two points $x, y \in \mathbb{M}$, taking x as the departure point, is denoted as γ_x^y, while a solution with prescribed departure point $x \in \mathbb{M}$ and velocity at departure $v \in T_x\mathbb{M}$ is denoted as $\gamma_{x,v}$. The length of a geodesic γ_x^y defines a distance $\mathscr{D}(x, y)$ between its endpoints.

To a geodesic arc $\gamma_{x,v} : [0, 1] \to \mathbb{M}$ is associated a map $\exp_x : T_x\mathbb{M} \to \mathbb{M}$ defined as $\exp_x(v) := \gamma_{x,v}(1)$. Such function is termed exponential map with pole $x \in \mathbb{M}$. The inverse function associated to the exponential map is termed logarithmic map. On a manifold $\mathbb{M}$ where it is defined an exponential map $\exp : T\mathbb{M} \to \mathbb{M}$, a logarithmic map is denoted as $\log : \mathbb{M}^2 \to T\mathbb{M}$. A logarithmic map satisfies the identity $\exp_x(\log_x y) = y$.

An orthogonal projector $\Pi_x : \mathbb{A} \to T_x\mathbb{M}$ is a function that maps an element of the ambient space to a tangent vector in a specific tangent space, namely $\Pi_x(a) \in T_x\mathbb{M} \Leftrightarrow a - \Pi_x(a) \in N_x\mathbb{M}$. A derivative of the orthogonal projector with respect to a curve is $\Pi^\bullet : \mathbb{A} \times T\mathbb{M} \to \mathbb{A}$, which is defined as

$$\Pi_\gamma^\bullet(a, \dot{\gamma}) := \frac{\mathrm{d}}{\mathrm{d}t} \Pi_\gamma(a), \tag{9}$$

where γ denotes any smooth curve.

Let us recall a property of the derivative $\Pi^\bullet$ that will turn useful in the following development. Whenever $\Pi^\bullet$ is applied to a pair of tangent vectors, it will return a normal vector, namely the restriction $\Pi^\bullet : T\mathbb{M} \times T\mathbb{M} \to N\mathbb{M}$.

Example 1. *Let us show the behavior of the derivative $\Pi^\bullet$ just described in the case of the manifold $\mathbb{SO}(n)$ endowed with its canonical metric. In this case $\Pi_R(A) = \frac{1}{2}(A - RA^\top R)$, therefore taken $W, V \in T_R\mathbb{SO}(n)$ and a curve $t \mapsto R(t)$ such that $\dot{R} = V$, it is seen that*

$$
\begin{aligned}
\Pi_R^\bullet(W, V) &= -\tfrac{1}{2}(VW^\top R + RW^\top V) \\
&= -\tfrac{1}{2}(VW^\top + RW^\top VR^\top)R \\
&= -\tfrac{1}{2}(VW^\top + WR^\top RV^\top)R \\
&= -\tfrac{1}{2}(VW^\top + WV^\top)R.
\end{aligned}
\tag{10}
$$

The matrix sum in parentheses is symmetric, hence $\Pi_R^\bullet(W, V)$ is a normal vector at R.

The Riemannian gradient of a smooth function $f : \mathbb{M} \to \mathbb{R}$ is denoted as $\mathrm{grad}_x f$ and may be written by the explicit formula $\mathrm{grad}_x f = G_x^{-1}(\Pi_x(\partial_x f))$, where $\partial_x f$ denotes the Gateaux derivative of f in $\mathbb{A}$. A golden calculation rule involving the notion of Riemannian gradient and that of Riemannian distance is $\mathrm{grad}_x \mathscr{D}^2(x, y) = -2 \log_x y$.

Parallel transport of a vector $v \in T_x\mathbb{M}$ along a geodesic joining points $x, y \in \mathbb{M}$ is indicated by $\mathbb{P}^{x \to y}(v)$. By this notation it is understood that x is the departure point. Its calculation is based on a Christoffel form of the second kind. It is convenient to define two Christoffel forms, namely the Christoffel form of the first kind $\bar{\Gamma} : (T\mathbb{M})^2 \to \mathbb{A}$ as

$$
\bar{\Gamma}_x := \tfrac{1}{2} G_x^\bullet \ + \ \text{a normal component},
\tag{11}
$$

and a Christoffel form of the second kind $\bar{\bar{\Gamma}} : (T\mathbb{M})^2 \to \mathbb{A}$ as $\bar{\bar{\Gamma}}_x := G_x^{-1}\bar{\Gamma}_x$. A property that involves parallel transport and manifold logarithm is $\mathbb{P}^{x \to y}(\log_x y) = -\log_y x$.

3. Covariant Derivative of a Vector Field

Covariant derivation provides a kind of derivative to evaluate the rate of change of a vector field along a given direction starting from a given point. Covariant derivation is a generalization to manifolds of the familiar notion of *directional derivative*.

The present section is organized as follows. Section 3.1 presents a brief review of the standard directional derivative and of its properties. Section 3.2 presents a coordinate-free conceptualization of covariant derivation along a curve, which is the main topic of the present section. Section 3.3 introduces the more general notion of covariant derivation of a vector field along a given direction and that of connection. Section 3.4 discusses again the notion of Riemannian Hessian operator and shows how it can be obtain by the covariant derivative of the Riemannian gradient. Section 3.5 discusses a definition of covariant derivation through axioms and shows that the introduced Levi-Civita connection meets such axioms. Section 3.6 retraces the notion of covariant derivation in coordinates. Section 3.7 introduces two new concepts, namely the commutator of two vector fields and the torsion field, which are indeed closely related. Section 3.8 illustrates and explains the consequence of a quite remarkable relationship between the Christoffel form of the second kind and the linearization of the projection operator. Section 3.9 discusses the notion of Lie derivative, a primeval attempt to define differentiation on tangent bundles.

3.1. Brief Review of Directional Derivative and of Its Properties

Let us briefly review the notion of direction derivative to better understand its manifold counterpart. Let us consider a smooth scalar multivariable function $f : \mathbb{R}^3 \to \mathbb{R}$. Given a point $p \in \mathbb{R}^3$, a directional derivative essentially measures the rate of change of the function f when moving away from p along a given direction. Two points come immediately to mind while endeavoring a definition for a directional derivative:

- Since f is a multivariable function, because its domain is $\mathbb{R}^3$, there are plenty of ways to 'move away' from the foot-point p. The rate of change of the function f depends on the direction towards which one moves away from point x, hence a direction

$v \in \mathbb{R}^3$ needs to be specified. For example, if one takes $f(x, y, z) := 0.1\, x^2 + 10\, y^2 + e^z$, the function f changes little along the x axis, changes more along the y axis, increases along the positive direction of the z axis and decreases along the negative direction of the z axis, not to mention the behavior corresponding to every possible combinations of the triple (x, y, z).

- Since f is arbitrary, its rate of change depends also on 'how far' one moves away from the foot-point p. Taking a large leap is, however, not very useful, as one may always replace the foot-point p with a new foot-point p' to know how the function behaves far from p. Therefore, it is understood that the rate of change is evaluated in close proximity of the point p.

Consequently, a possible definition of directional derivative is

$$D_p f(v) := \lim_{h \to 0} \frac{f(p + h\,v) - f(p)}{h} \in \mathbb{R} \text{ for } p \in \mathbb{R}^3,\ v \in \mathbb{R}^3. \tag{12}$$

From multivariable calculus it is known that

$$D_p f(v) = \left. \frac{\mathrm{d}f(p + t\,v)}{\mathrm{d}t} \right|_{t=0} = \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y + \frac{\partial f}{\partial z} v_z = (\partial_p f)^\top v, \tag{13}$$

where $v = (v_x, v_y, v_z)$ and $\partial_p f$ denotes the gradient (or Jacobian, or Gateaux derivative) of the function f with respect to its three scalar arguments, namely $\partial_p f := \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$. The rightmost quantity is, according to the general definition (6), the differential $\mathrm{d}_p f(v)$ of the function f, at point p, applied to the vector v. In this case, however, since we are working in $\mathbb{R}^3$, the 'vectorial' nature of v is not immediately apparent.

Let us now generalize this notion by taking a smooth function $f : \mathbb{R}^m \to \mathbb{R}^n$, which is an array-type multivariable function, which we can call a *multivariable vector field*. The function f assigns an array to any point of its domain. Therefore $f(p)$ may be though of as a vector field (once again, its vectorial nature as understood in manifold calculus in not apparent). Like before, a notion of 'rate of change' of a vector field along a given direction may be expressed as

$$D_p f(v) := \lim_{h \to 0} \frac{f(p + h\,v) - f(p)}{h} \in \mathbb{R}^n \text{ for } p \in \mathbb{R}^m,\ v \in \mathbb{R}^m. \tag{14}$$

Although the definition just given is formally identical to (12), the space where the two directional derivatives live are different. From multivariable calculus we get:

$$(D_p f(v))_i = \left(\left. \frac{\mathrm{d}f(p + t\,v)}{\mathrm{d}t} \right|_{t=0} \right)_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial p_j} v_j, \tag{15}$$

where p_i denotes the ith component of the m-dimensional array p, v_i denotes the ith component of the m-dimensional array v and f_i denotes the ith component of the array-valued function f (with respect to the canonical basis of the domain $\mathbb{R}^m$ and of the codomain space $\mathbb{R}^n$). Hence, in this generalized scenario, $D_p f(v) = (\partial_p f) v$, where $\partial_p f$ denotes the Jacobian matrix of the transformation $p \mapsto f(p)$. In this case, the rate of change itself is multivariable, hence this directional derivative is related to the notion of differential in the sense that $(D_p f(v))_i = \mathrm{d}_p f_i(v)$.

To complete this brief survey of directional derivation, let us emphasize a few properties that the directional derivative $D_p f$ enjoys:

- ($\mathbf{P_1}$) Assume that the direction v along which the rate of change is sought is given by the superposition of two partial directions, namely $v = \alpha u + \beta w$, with α, β being two scalar fields, namely $\alpha = \alpha(p)$ and $\beta = \beta(p)$. How the rate of change $D_p f(v)$ relates to rates of change $D_p f(u)$ and $D_p f(w)$ along partial directions is easily understood:

$$D_p f(\alpha u + \beta w) = (\partial_p f)(\alpha u + \beta w) = \alpha(\partial_p f)u + \beta(\partial_p f)w = \alpha D_p f(u) + \beta D_p f(w). \tag{16}$$

- **(P$_2$)** Assume the field f be given as the superposition of two concurrent fields, namely $f = b + g$. The way the rate of change $D_p f(v)$ relates to rates of change $D_p b(v)$ and $D_p g(v)$ is also easily understood:

$$D_p f(v) = D_p b(v) + D_p g(v). \tag{17}$$

- **(P$_3$)** In addition, assume the field f is given as the dilation of another field, namely $f = \alpha g$, with α being a scalar field. How the rate of change $D_p f(v)$ relates to the rate of change $D_p g(v)$ is readily understood:

$$D_p f(v) = g \, D_p \alpha(v) + \alpha \, D_p g(v) = g \, d_p \alpha(v) + \alpha \, D_p g(v) \tag{18}$$

by the rule of partial derivation of compound functions.

Having recalled the notion of directional derivative of a vector field in $\mathbb{R}^m$, we are now going to examine and discuss its extension to manifolds.

3.2. Coordinate-Free Conceptualization of Covariant Derivation along a Curve

The notion of covariant derivation may be conceptualized in different manners. We recall here two relevant ways:

- **Top-down from parallel transport**: Assuming that a notion of parallel transport is available, which enables us to compare the values of a vector field at different points of a manifold, covariant derivation stems quite naturally by adapting the classical definition of directional derivative (14). In this sense, covariant derivation represents a 'local' version of parallel transport.
- **Direct axiomatization**: In this instance, covariant derivative is defined to be an operator that meets certain requirements, which basically retrace properties P$_1$, P$_2$, P$_3$ noted in Section 3.1. Such definition is quite general and leads to a family of covariant derivatives. Moreover, such definition is given independently from that of parallel transport, which may, in turn, be defined as the solution of a covariant-derivative-based differential equation on the tangent bundle $T\mathbb{M}$. In this sense, parallel transport represents a 'global' version of covariant derivation.

In the present tutorial paper, we are going to follow the first way of defining covariant derivation (based on parallel transport), not only because one such flow of discussion would be coherent with the general setting of the present series of tutorials, but also because the top-down conceptualization is deemed to appear more clear to control engineers and applied scientist and well-suited to develop numerical methods to simulate the behavior of second-order dynamical systems on manifold.

As a first notion to be defined, we are going to deploy *covariant derivation along a curve*. Namely, we assume that a vector field $w \in \Gamma(\mathbb{M})$ is well-defined and regular along a smooth curve $\gamma : [a, \, b] \to \mathbb{M}$ and we need to evaluate the rate of change of the vector field $\gamma \mapsto w_\gamma$ along the curve itself. Such rate of change will be denoted as $\nabla_t^\gamma w$, where $t \in [a, \, b]$ denotes the value of the parameter that individuates univocally a point over the curve γ. Such rate of change is precisely the covariant derivative of the vector field w along the curve γ at the point t, and is defined as

$$\nabla_t^\gamma w := \lim_{h \to 0} \frac{\mathbb{P}^{\gamma(t+h) \to \gamma(t)} w_{\gamma(t+h)} - w_{\gamma(t)}}{h} = \frac{\mathrm{d}}{\mathrm{d}\tau} \mathbb{P}^{\gamma(\tau) \to \gamma(t)} w_{\gamma(\tau)} \Big|_{\tau = t}. \tag{19}$$

Such definition appears as a natural extension of the directional derivative (14) of a vector field, properly corrected by parallel transport to align vectors belonging to different tangent spaces. It should be noticed that the curve γ does not necessarily need to be a geodesic. In fact, in system and control theory, it likely represents a trajectory generated

by a dynamical system over its state manifold, which only occasionally coincides with a geodesic arc.

In order to derive an explicit expression of the covariant derivative, let us evaluate the time-derivative of the function $\mathscr{Q} : [a,\, b] \to \mathbb{R}$ defined as follows:

$$\mathscr{Q}(t) := \langle w_{\gamma(t)}, \dot{\gamma}(t) \rangle_{\gamma(t)}. \tag{20}$$

By definition of derivative with respect to the scalar parameter t, we have

$$\dot{\mathscr{Q}}(t) = \lim_{h \to 0} \frac{\langle w_{\gamma(t+h)}, \dot{\gamma}(t+h) \rangle_{\gamma(t+h)} - \langle w_{\gamma(t)}, \dot{\gamma}(t) \rangle_{\gamma(t)}}{h}. \tag{21}$$

By the property of conformal isometry of parallel transport (with respect to the inner product), it follows that

$$\langle w_{\gamma(t+h)}, \dot{\gamma}(t+h) \rangle_{\gamma(t+h)} = \langle \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(w_{\gamma(t+h)}), \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(\dot{\gamma}(t+h)) \rangle_{\gamma(t)}, \tag{22}$$

hence the derivative $\dot{\mathscr{Q}}$ may be rewritten as

$$\begin{aligned}
\dot{\mathscr{Q}}(t) &= \lim_{h \to 0} \frac{\langle \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(w_{\gamma(t+h)}), \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(\dot{\gamma}(t+h)) \rangle_{\gamma(t)} - \langle w_{\gamma(t)}, \dot{\gamma}(t) \rangle_{\gamma(t)}}{h} \\
&= \lim_{h \to 0} \frac{\langle \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(w_{\gamma(t+h)}) - w_{\gamma(t)} + w_{\gamma(t)}, \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(\dot{\gamma}(t+h)) - \dot{\gamma}(t) + \dot{\gamma}(t) \rangle_{\gamma(t)} - \langle w_{\gamma(t)}, \dot{\gamma}(t) \rangle_{\gamma(t)}}{h} \\
&= \lim_{h \to 0} \frac{\langle \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(w_{\gamma(t+h)}) - w_{\gamma(t)}, \dot{\gamma}(t) \rangle_{\gamma(t)}}{h} \\
&\quad + \lim_{h \to 0} \frac{\langle w_{\gamma(t)}, \mathbb{P}^{\gamma(t+h) \to \gamma(t)}(\dot{\gamma}(t+h)) - \dot{\gamma}(t) \rangle_{\gamma(t)}}{h} \\
&= \langle \nabla_t^{\gamma} w, \dot{\gamma}(t) \rangle_{\gamma(t)} + \langle w_{\gamma(t)}, \nabla_t^{\gamma} \dot{\gamma} \rangle_{\gamma(t)}.
\end{aligned} \tag{23}$$

On the basis of the above development, it is possible to write an explicit expression for the covariant derivative of a vector field along a curve based on the (derivative of) the metric kernel. Let us rewrite the function $\mathscr{Q}$ as $\mathscr{Q}(t) = \langle w_{\gamma(t)}, G_{\gamma(t)}(\dot{\gamma}(t)) \rangle^{\mathbb{A}}$. Then, it holds that

$$\dot{\mathscr{Q}} = \langle \dot{w}_{\gamma}, G_{\gamma}(\dot{\gamma}) \rangle^{\mathbb{A}} + \langle w_{\gamma}, G_{\gamma}^{\bullet}(\dot{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}} + \langle w_{\gamma}, G_{\gamma}(\ddot{\gamma}) \rangle^{\mathbb{A}}, \tag{24}$$

where $\dot{w}_{\gamma}$ denotes, for brevity, the total naïve derivative $\frac{\mathrm{d}}{\mathrm{d}t}(w_{\gamma(t)})$. Notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}(w_{\gamma}) = (D_{\gamma}w)(\dot{\gamma}) = (\partial_{\gamma}w)(\dot{\gamma}), \tag{25}$$

where $\partial_x w$ denotes the usual Gateaux derivative and appears as a linear operator (apparently, such operation is defined in the ambient space $\mathbb{A}$ rather than in $\mathbb{M}$). Let us rewrite the above expression by introducing a simple trick:

$$\dot{\mathscr{Q}} = \underbrace{\langle G_{\gamma}(\dot{\gamma}), \dot{w}_{\gamma} \rangle^{\mathbb{A}}}_{(A)} + \tfrac{1}{2} \underbrace{\langle w_{\gamma}, G_{\gamma}^{\bullet}(\dot{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}}}_{(B)} + \underbrace{\langle w_{\gamma}, G_{\gamma}(\ddot{\gamma}) \rangle^{\mathbb{A}}}_{(C)} + \tfrac{1}{2} \underbrace{\langle w_{\gamma}, G_{\gamma}^{\bullet}(\dot{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}}}_{(D)}. \tag{26}$$

Let us examine each term separately. The term (A) in (26) may be rewritten as

$$\langle G_{\gamma}(\dot{\gamma}), \dot{w}_{\gamma} \rangle^{\mathbb{A}} = \langle G_{\gamma}(\dot{\gamma}), \Pi_{\gamma}(\dot{w}_{\gamma}) \rangle^{\mathbb{A}} \tag{27}$$

by definition of orthogonal projection.

The term (B) may be rewritten as $\langle w_{\gamma}, G_{\gamma}^{\bullet}(\dot{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}} = \langle \dot{\gamma}, G_{\gamma}^{\bullet}(w_{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}}$ thanks to the commutativity of the derivative $G^{\bullet}$ with respect to the ambient inner product. Moreover, by adding an identity $G_{\gamma} \circ G_{\gamma}^{-1}$ gives $\langle \dot{\gamma}, G_{\gamma}^{\bullet}(w_{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}} = \langle G_{\gamma} G_{\gamma}^{-1} \dot{\gamma}, G_{\gamma}^{\bullet}(w_{\gamma}, \dot{\gamma}) \rangle^{\mathbb{A}}$, where cumbersome parentheses were (and will be) omitted for brevity. Invoking again the

property of the metric kernel to be self-adjoint and the definition of orthogonal projection leads to the identity

$$\langle w_\gamma, G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}} = \langle G_\gamma(\dot\gamma), \Pi_\gamma G_\gamma^{-1} G_\gamma^\bullet(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}}, \tag{28}$$

where we have used quite liberally the 'extendability' property of the metric kernel to the ambient space.

The term (C) may be rewritten as $\langle w_\gamma, G_\gamma(\ddot\gamma)\rangle^{\mathbb{A}} = \langle G_\gamma(w_\gamma), \ddot\gamma\rangle^{\mathbb{A}}$ and, introducing an orthogonal projector, we get $\langle G_\gamma(w_\gamma), \ddot\gamma\rangle^{\mathbb{A}} = \langle G_\gamma(w_\gamma), \Pi_\gamma(\ddot\gamma)\rangle^{\mathbb{A}}$, hence

$$\langle w_\gamma, G_\gamma(\ddot\gamma)\rangle^{\mathbb{A}} = \langle G_\gamma(w_\gamma), \Pi_\gamma(\ddot\gamma)\rangle^{\mathbb{A}}. \tag{29}$$

It is now worth noticing that the term (D) may be rewritten equivalently as $\langle w_\gamma, G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}} = \langle w_\gamma, G_\gamma G_\gamma^{-1} G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}}$ and, in addition, observing that $\langle w_\gamma, G_\gamma G_\gamma^{-1} G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}} = \langle G_\gamma(w_\gamma), G_\gamma^{-1} G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}}$, we may conclude that

$$\langle w_\gamma, G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}} = \langle G_\gamma(w_\gamma), \Pi_\gamma G_\gamma^{-1} G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}}. \tag{30}$$

Summing up the four terms in (26) written in the indicated ways, we obtain the expression

$$\dot{\mathscr{Q}} = \langle G_\gamma(\dot\gamma), \Pi_\gamma(\dot w_\gamma) + \tfrac{1}{2}\Pi_\gamma G_\gamma^{-1} G_\gamma^\bullet(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}} + \\ \langle G_\gamma(w_\gamma), \Pi_\gamma(\ddot\gamma) + \tfrac{1}{2}\Pi_\gamma G_\gamma^{-1} G_\gamma^\bullet(\dot\gamma, \dot\gamma)\rangle^{\mathbb{A}}. \tag{31}$$

Comparing the above expression with the relationship (23) and recalling that the vector field w_γ is arbitrary, one gets the two identities

$$\nabla_t^\gamma w = \Pi_\gamma\{\tfrac{d}{dt}(w_\gamma) + \tfrac{1}{2}G_\gamma^{-1}(G_\gamma^\bullet(w_\gamma, \dot\gamma))\}, \tag{32}$$

$$\nabla_t^\gamma \dot\gamma = \Pi_\gamma\{\ddot\gamma + \tfrac{1}{2}G_\gamma^{-1}(G_\gamma^\bullet(\dot\gamma, \dot\gamma))\}. \tag{33}$$

The two above expressions are coherent, in fact, setting $w_\gamma := \dot\gamma$ in the relationship (32) yields the relationship (33).

The relationship (32) may be rewritten through the Christoffel form of the first kind. In fact from the expression (11) it follows that the term (B) in (26) may be rewritten as $\langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}} = 2\langle \dot\gamma, \bar\Gamma_\gamma(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}} = 2\langle G_\gamma(\dot\gamma), G_\gamma^{-1}\bar\Gamma_\gamma(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}} = 2\langle G_\gamma(\dot\gamma), \Pi_\gamma G_\gamma^{-1}\bar\Gamma_\gamma(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}} = 2\langle G_\gamma(\dot\gamma), \Pi_\gamma\bar{\bar\Gamma}_\gamma(w_\gamma, \dot\gamma)\rangle^{\mathbb{A}}$. As a consequence, the covariant derivative (32) may be rewritten equivalently as

$$\nabla_t^\gamma w = \Pi_\gamma\{\tfrac{d}{dt}(w_\gamma) + \bar{\bar\Gamma}_\gamma(w_\gamma, \dot\gamma)\}, \tag{34}$$

a classical expression in manifold calculus. In terms of Gateaux derivative of the vector field w, we may also write

$$\nabla_t^\gamma w = \Pi_\gamma\{(\partial_\gamma w_\gamma)(\dot\gamma) + \bar{\bar\Gamma}_\gamma(w_\gamma, \dot\gamma)\}, \tag{35}$$

which evidences how $\nabla_t^\gamma w$ is linear in $\dot\gamma$.

Two special cases of the above relationship are particularly worth examining separately:

- **Case of normal Christoffel form**. Whenever the manifold $\mathbb{M}$ is endowed with a normal Christoffel form $\bar{\bar\Gamma}_x$, namely whenever $\Pi_x\bar{\bar\Gamma}_x \equiv 0$, the general expression of the covariant derivative may be simplified as

$$\nabla_t^\gamma w = \Pi_\gamma\{\tfrac{d}{dt}(w_\gamma)\}, \tag{36}$$

  namely, the covariant derivative of a vector field coincides with an orthogonal projection of the naïve derivative of the vector field.

- **Case of geodesics**. When $w_\gamma = \dot\gamma$, namely the vector field under derivation coincides with the velocity field of a curve, it holds that $\nabla_t^\gamma \ddot\gamma = \Pi_\gamma\{\ddot\gamma + \bar{\bar\Gamma}_\gamma(\dot\gamma, \dot\gamma)\}$. If γ is a

geodesic curve, since $\dot{\gamma} + \bar{\bar{\Gamma}}_{\gamma}(\dot{\gamma}, \dot{\gamma}) = 0$ in every point of a geodesic, then it turns out that, on every geodesic line,

$$\nabla_t^{\gamma} \dot{\gamma} = 0. \tag{37}$$

In fact, in a 'direct axiomatization' setting, the above relationship is taken as one defining a geodesic!

We shall revise the expression (34) in Section 3.8 where we shall show that projection is not strictly necessary, although taking out the projection operator does not necessarily simplify the related calculations!

From the above calculations one might infer that, for all practical purposes, the operators G_x and Π_x commute, namely $G_x \Pi_x = \Pi_x G_x$.

Example 2. *In all instances in which $G_x \equiv \mathrm{id}_x$, as in the case of the unit hypersphere endowed with its canonical metric, the property $G_x \Pi_x = \Pi_x G_x$ is trivially verified.*

Let us consider the case of the manifold $\mathbb{S}^+(n)$ endowed with its canonical metric, for which $G_P(W) = P^{-1} W P^{-1}$ and $\Pi(A) = \frac{1}{2}(A^{\top} + A)$ (uniform). In this case

$$\begin{aligned}
\Pi\, G_P(A) &= \tfrac{1}{2}((P^{-1} A P^{-1})^{\top} + P^{-1} A P^{-1}) \\
&= \tfrac{1}{2}(P^{-\top} A^{\top} P^{-\top} + P^{-1} A P^{-1}) \\
&= \tfrac{1}{2}(P^{-1} A^{\top} P^{-1} + P^{-1} A P^{-1}) \\
&= P^{-1} \tfrac{A^{\top} + A}{2} P^{-1} \\
&= G_P \Pi(A).
\end{aligned} \tag{38}$$

Let us further consider the case of the Stiefel manifold $\mathrm{St}(n, p)$ endowed with its canonical metric, where $G_X(W) = (I_n - \frac{1}{2} X X^{\top}) W$ and $\Pi_X(A) = A - \frac{1}{2} X(X^{\top} A + A^{\top} X)$. In this case, it holds

$$\begin{aligned}
\Pi_X G_X(A) &= (I_n - \tfrac{1}{2} X X^{\top}) A - \tfrac{1}{2} X(X^{\top}(I_n - \tfrac{1}{2} X X^{\top}) A + A^{\top}(I_n - \tfrac{1}{2} X X^{\top}) X) \\
&= A - \tfrac{3}{4} X X^{\top} A - \tfrac{1}{4} X A^{\top} X, \\
G_X \Pi_X(A) &= (I_n - \tfrac{1}{2} X X^{\top})(A - \tfrac{1}{2} X(X^{\top} A + A^{\top} X)) \\
&= A - \tfrac{3}{4} X X^{\top} A - \tfrac{1}{4} X A^{\top} X,
\end{aligned} \tag{39}$$

hence the property of commutativity of the two operators is verified once again.

In fact, the commutation property just exemplified may be proved quite easily by observing that, for every $a \in \mathbb{A}$ and $v \in T_x \mathbb{M}$, the following equality chain holds:

$$\begin{aligned}
\langle G_x \Pi_x \{a\}, v \rangle_x &= \langle G_x \Pi_x \{a\}, G_x(v) \rangle^{\mathbb{A}} = \langle \Pi_x \{a\}, G_x^2(v) \rangle^{\mathbb{A}} = \langle a, G_x^2(v) \rangle^{\mathbb{A}} \\
&= \langle G_x(a), G_x(v) \rangle^{\mathbb{A}} = \langle \Pi_x G_x(a), G_x(v) \rangle^{\mathbb{A}} = \langle \Pi_x G_x(a), v \rangle_x,
\end{aligned} \tag{40}$$

where we used the shortened notation G_x^2 for $G_x \circ G_x$. The above equality chain implies that $G \Pi = \Pi G$.

Let us consider a few examples of calculation of covariant derivative.

Example 3. *As an example of calculation of covariant derivative, let us consider the case of the unit hypersphere $\mathbb{S}^{n-1}$ endowed with its canonical metric. The orthogonal projector reads $\Pi_x(a) = (I_n - x x^{\top}) a$. According to the relationship (36), the covariant derivative would read*

$$\nabla_t^{\gamma} w = (I_n - \gamma \gamma^{\top}) \tfrac{\mathrm{d}}{\mathrm{d}t}(w_{\gamma}). \tag{41}$$

Let us verify the same result by applying the definition (19). Let us recall the expression of the parallel transport operator on the hyper-sphere $\mathbb{S}^{n-1}$ that reads $\mathbb{P}^{x \to y}(u) = \left(I_n - \frac{(x+y) y^{\top}}{1 + x^{\top} y} \right) u$. By direct calculation we get

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \mathbb{P}^{\gamma(\tau)\to\gamma(t)} w_{\gamma(\tau)} = \frac{\mathrm{d}}{\mathrm{d}\tau} \left(w_{\gamma(\tau)} - \frac{(\gamma(\tau)+\gamma(t))\gamma^\top(t)}{1+\gamma^\top(\tau)\gamma(t)} w_{\gamma(\tau)} \right)$$

$$= \dot{w}_{\gamma(\tau)}$$

$$- \frac{\dot{\gamma}(\tau)\gamma^\top(t)(1+\gamma^\top(\tau)\gamma(t)) - (\gamma(\tau)+\gamma(t))\gamma^\top(t)\dot{\gamma}^\top(\tau)\gamma(t)}{\{1+\gamma^\top(\tau)\gamma(t)\}^2} w_{\gamma(\tau)} \qquad (42)$$

$$- \frac{(\gamma(\tau)+\gamma(t))\gamma^\top(t)}{1+\gamma^\top(\tau)\gamma(t)} \dot{w}_{\gamma(\tau)}.$$

Setting $\tau = t$ leads to

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \mathbb{P}^{\gamma(\tau)\to\gamma(t)} w_{\gamma(\tau)} \Big|_{\tau=t} = \dot{w}_{\gamma(t)}$$

$$- \frac{2\dot{\gamma}(t)\gamma^\top(t) - 2\gamma(t)\gamma^\top(t)[\dot{\gamma}^\top(t)\gamma(t)]}{4} w_{\gamma(t)} - \frac{2\gamma(t)\gamma^\top(t)}{2}\dot{w}_{\gamma(t)} \quad (43)$$

$$= \dot{w}_{\gamma(t)} - \tfrac{1}{2}\dot{\gamma}(t)(\gamma^\top(t)w_{\gamma(t)}) - \gamma(t)\gamma^\top(t)\dot{w}_{\gamma(t)}$$

$$= (I_n - \gamma(t)\gamma^\top(t))\dot{w}_{\gamma(t)}.$$

Hence the two results coincide, as expected.

Let us now compute the covariant derivative of a vector field in $\Gamma(\mathbb{S}^+(n))$. Let us recall that, for the manifold of symmetric, positive-definite matrices, we found (cfr. [13]) that $\bar{\bar{\Gamma}}_P(V, W) = -\frac{1}{2}(VP^{-1}W + WP^{-1}V)$. It is straightforward to verify that, for every $P \in \mathbb{S}^+(n)$, $V, W \in T_P\mathbb{S}^+(n)$ it holds that $\bar{\bar{\Gamma}}_P(V, W) \in T_P\mathbb{S}^+(n)$, as well as that, for every vector field $W \in \Gamma(\mathbb{S}^+(n))$ and every smooth trajectory $t \mapsto P(t)$, it holds that the naïve derivative $\frac{\mathrm{d}}{\mathrm{d}t}W_{P(t)} \in T_{P(t)}\mathbb{S}^+(n)$. As a consequence, in the present case the general relationship (34) simplifies into

$$\nabla_t^P W = \frac{\mathrm{d}}{\mathrm{d}t}W_P + \bar{\bar{\Gamma}}_P(W_P, \dot{P}), \qquad (44)$$

which, by the expression of the Christoffel form, reads

$$\nabla_t^P W = \frac{\mathrm{d}}{\mathrm{d}t}W_P - \tfrac{1}{2}(W_P P^{-1}\dot{P} + \dot{P}P^{-1}W_P), \qquad (45)$$

To end with this series of examples, let us compute the covariant derivative of a vector field in $\Gamma(\mathbb{SO}(n))$. Let us recall that, whenever the manifold $\mathbb{SO}(n)$ is endowed with its canonical metric, it holds $\Pi_R(A) = \frac{1}{2}A - \frac{1}{2}RA^\top R$, while the Christoffel form of the second kind takes the expression $\bar{\bar{\Gamma}}_R(V, W) = \frac{1}{2}R(V^\top W + W^\top V)$. Indeed, the exact expression of the Christoffel form, in this case, is irrelevant, since it maps $T\mathbb{M} \times T\mathbb{M}$ to the normal bundle $N\mathbb{M}$. Therefore, over a trajectory $t \mapsto R(t)$, the covariant derivative of a vector field $R \mapsto W_R$ reads

$$\nabla_t^R W = \frac{\mathrm{d}}{\mathrm{d}t}W_R - R\left(\frac{\mathrm{d}}{\mathrm{d}t}W_R\right)^\top R. \qquad (46)$$

It is a useful exercise to verify by direct calculations from the definition that the above expression is correct.

3.3. Covariant Derivation along a Direction, Connections

Let $\nabla_t^\gamma w$ denote the covariant derivative of a vector field $x \mapsto w_x$ at a point t of a curve γ. In some cases, one might want to indicate explicitly the direction along which the derivative is taken, which coincides with $\dot{\gamma}(t)$. Referring the foot-point and the tangent direction at $t = 0$, in these cases one may set $v := \dot{\gamma}(0)$ and denote the covariant derivative by the shortened notation

$$\nabla_v w := \nabla_t^\gamma w\big|_{t=0}. \qquad (47)$$

In this case, the covariant derivative does not depend on the curve γ, except for $\gamma(0)$ and $\dot{\gamma}(0)$. (If it is necessary to specify the foot-point $x \in \mathbb{M}$ where all these computations take place, the elongated notation $(\nabla_v w)_x$ may be used). For example, the covariant

derivative of a velocity field with respect to itself, namely the *covariant acceleration*, would read $\nabla_{\dot{\gamma}}\dot{\gamma}$.

An extension of the covariant derivative of a vector field along a given direction leads to the notion of *connection* $\nabla : \Gamma(\mathbb{M}) \times \Gamma(\mathbb{M}) \to \Gamma(\mathbb{M})$. Let us consider a vector field $w \in \Gamma(\mathbb{M})$ that is to be derivated and a further vector field $v \in \Gamma(\mathbb{M})$ that associates to each point of the manifold (or a region of the manifold) a direction along which derivate. In other terms, at a point $x \in \mathbb{M}$, v_x defines the direction along which the rate of change of the vector field w is sought. The corresponding covariant derivative, denoted by $x \mapsto \nabla_{v_x} w_x$ or $(\nabla_v w)_x$, extends the notion of covariant derivation of a vector field along a curve and is based on the assumption that the vector field w is defined in every point of $\mathbb{M}$ (or within a region of the manifold). The formal definition goes like

$$x \mapsto (\nabla_v w)_x := \lim_{h \to 0} \frac{\mathbb{P}^{\exp_x(h\,v_x) \to x} w_{\exp_x(h\,v_x)} - w_x}{h}. \tag{48}$$

The connection described in the present tutorial is commonly referred to as *Levi-Civita connection*.

A *metric connection* is a covariant derivative that 'agrees' or is 'coordinated' with an inner product in the sense that the covariant derivative of an inner product of two vector fields depends only on the rate of change of such vector fields and does not depend on the rate of change of the inner product itself, namely, given two vector fields $v, w \in \Gamma(\mathbb{M})$ and a tangent direction $u \in T_x\mathbb{M}$, it holds that

$$\nabla_u \langle v_x, w_x \rangle_x = \langle (\nabla_u v)_x, w_x \rangle_x + \langle v_x, (\nabla_u w)_x \rangle_x. \tag{49}$$

The definition (19) ensures that the resulting connection is metric. Notice, in fact, that $x \mapsto \langle u_x, w_x \rangle_x$ is a *scalar field*, a special case of vector field, that may be subjected to covariant derivation. In addition, notice that in a large number of references, the directional derivative $\nabla_v \langle u_x, w_x \rangle_x$ is denoted simply as $v(\langle u, w \rangle)$, leaving understood that a vector field acts as a directional derivative.

Example 4. *We can see this even in the case of a flat manifold $\mathbb{R}^3$ endowed with an inner product $\langle u_x, w_x \rangle_x$. In this case, we have that the directional derivative of the inner product of two vector fields along a direction v is*

$$\nabla_v \langle u_x, w_x \rangle_x := \lim_{h \to 0} \frac{\langle u_{x+hv}, w_{x+hv} \rangle_{x+hv} - \langle u_x, w_x \rangle_x}{h}. \tag{50}$$

Adding zero-sum terms, and rearranging them, gives:

$$\begin{aligned}
\nabla_v \langle u_x, w_x \rangle_x := &\lim_{h \to 0} \frac{\langle u_{x+hv}, w_{x+hv} \rangle_{x+hv} - \langle v_{x+hv}, w_x \rangle_{x+hv}}{h} + \\
&\lim_{h \to 0} \frac{\langle u_{x+hv}, w_{x+hv} \rangle_{x+hv} - \langle v_x, w_{x+hv} \rangle_{x+hv}}{h} + \\
&\lim_{h \to 0} \frac{\langle u_x, w_x \rangle_{x+hv} - \langle u_x, w_x \rangle_x}{h} + \\
&\lim_{h \to 0} \frac{\langle u_{x+hv}, w_x \rangle_{x+hv} - \langle u_{x+hv}, w_{x+hv} \rangle_{x+hv}}{h} + \\
&\lim_{h \to 0} \frac{\langle u_x, w_{x+hv} \rangle_{x+hv} - \langle u_x, w_x \rangle_{x+hv}}{h}.
\end{aligned} \tag{51}$$

By the linearity and continuity of the inner product one then gets

$$\begin{aligned}
\nabla_v \langle u_x, w_x \rangle_x = &\langle u_x, (\nabla_v w)_x \rangle_x + \langle (\nabla_v u)_x, w_x \rangle_x - \langle u_x, (\nabla_v w)_x \rangle_x + \\
&\langle u_x, (\nabla_v w)_x \rangle_x + \lim_{h \to 0} \frac{\langle u_x, w_x \rangle_{x+hv} - \langle u_x, w_x \rangle_x}{h}.
\end{aligned} \tag{52}$$

In conclusion, one gets

$$\nabla_v \langle u_x, w_x \rangle_x = \langle u_x, (\nabla_v w)_x \rangle_x + \langle (\nabla_v u)_x, w_x \rangle_x + \lim_{h \to 0} \frac{\langle u_x, w_x \rangle_{x+hv} - \langle u_x, w_x \rangle_x}{h}. \tag{53}$$

The first two addenda depend only on the directional derivatives of the vector fields, while the third term depends on the rate of change of the inner product along the line $x + hv$. Only when the last term is identically zero the directional derivative is said to be metric, *in which case $\nabla_v \langle u_x, w_x \rangle_x = \langle u_x, (\nabla_v w)_x \rangle_x + \langle (\nabla_v u)_x, w_x \rangle_x$.*

3.4. Relationship of Gradient to Hessian by Covariant Derivation

A noticeable result from manifold calculus is that Riemannian Hessian, covariant derivation and Riemannian gradient are related to one another in a neat way, namely

$$H_\gamma f(\dot\gamma) = \nabla_t^\gamma (\mathrm{grad}_\gamma f). \tag{54}$$

The gradient builds a vector field over a curve based on a scalar function and the Hessian represents the rate of change of such vector field along the curve. In order to prove such neat result, it is convenient to re-derive the expression of the Riemannian gradient in a different way compared to the one derived in [13].

Let us recall that, given a scalar function $f : \mathbb{M} \to \mathbb{R}$ and a smooth curve γ with image in $\mathbb{M}$, it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma) = \langle \mathrm{grad}_\gamma f, \dot\gamma \rangle_\gamma = \langle \mathrm{grad}_\gamma f, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}}, \tag{55}$$

where $\mathbb{A}$ denotes, as usual, the ambient space that the manifold $\mathbb{M}$ is embedded in and $\langle \star, \star \rangle^{\mathbb{A}}$ denotes an Euclidean scalar product that the ambient space is endowed with. Deriving with respect to t the rightmost expression gives

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma) = \langle \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \langle \mathrm{grad}_\gamma f, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \rangle^{\mathbb{A}} + \langle \mathrm{grad}_\gamma f, G_\gamma(\ddot\gamma) \rangle^{\mathbb{A}}. \tag{56}$$

Let us examine separately the above three addenda. The first addendum may be rewritten as

$$\langle \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = \langle \Pi_\gamma \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = \langle \Pi_\gamma \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f, \dot\gamma \rangle_\gamma. \tag{57}$$

The second addendum may be rewritten as

$$\begin{aligned}
\langle \mathrm{grad}_\gamma f, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \rangle^{\mathbb{A}} &= \langle G_\gamma^\bullet(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle^{\mathbb{A}} = \langle 2\bar\Gamma_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle^{\mathbb{A}} = \\
&2\langle G_\gamma^{-1}\bar\Gamma_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = 2\langle \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = \\
&2\langle \Pi_\gamma \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = 2\langle \Pi_\gamma \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle_\gamma.
\end{aligned} \tag{58}$$

Since the expression (56) is path-dependent, we choose the curve γ to be a geodesic line, hence the third addendum takes the equivalent expressions

$$\begin{aligned}
\langle \mathrm{grad}_\gamma f, G_\gamma(\ddot\gamma) \rangle^{\mathbb{A}} &= \langle \mathrm{grad}_\gamma f, -\bar\Gamma_\gamma(\dot\gamma, \dot\gamma) \rangle^{\mathbb{A}} = -\langle \bar\Gamma_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle^{\mathbb{A}} = \\
&-\langle G_\gamma^{-1}\bar\Gamma_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = -\langle \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = \\
&-\langle \Pi_\gamma \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} = -\langle \Pi_\gamma \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle_\gamma.
\end{aligned} \tag{59}$$

Gluing pieces together, one obtains

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma) = \langle \Pi_\gamma \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f + \Pi_\gamma \bar{\bar\Gamma}_\gamma(\mathrm{grad}_\gamma f, \dot\gamma), \dot\gamma \rangle_\gamma. \tag{60}$$

By definition, the Hessian of a scalar function at a point $x \in \mathbb{M}$, along a given tangent direction $v \in T_x \mathbb{M}$, is such that

$$\left. \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma) \right|_{t=0} = \langle H_x f(v), v \rangle_x, \tag{61}$$

where $\gamma(0) = x$ and $\dot{\gamma}(0) = v$, hence we may conclude that

$$H_x f(v) = \Pi_x \left\{ \left. \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma f \right|_{t=0} + \bar{\bar{\Gamma}}_x (\mathrm{grad}_x f, v) \right\}. \tag{62}$$

Let us examine two examples of calculation of Hessian.

Example 5. *As a first example, let us consider the function $f(x) := \frac{1}{2} \mathscr{D}^2(x, r)$, where $\mathscr{D} : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ denotes a Riemannian distance and $r \in \mathbb{M}$ denotes a fixed reference point (This function is sometimes referred to as Synge's world function in relativistic physics [14]). As we have proven, in this case $\mathrm{grad}_x f = - \log_x r$, where $\log_x : \mathbb{M} \to T_x \mathbb{M}$ denotes a logarithmic map. In this example, therefore, it turns out that*

$$\frac{1}{2} (H_x \mathscr{D}^2(x, r))(v) = -\Pi_x \left\{ \left. \frac{\mathrm{d}}{\mathrm{d}t} \log_\gamma(r) \right|_{t=0} + \bar{\bar{\Gamma}}_x (\log_x r, v) \right\}. \tag{63}$$

The papers [15,16] further elaborate on the notion of Hessian of the Riemannian squared distance.

As a second example, let us consider the Hessian of Oja's function (or Rayleigh quotient) $f(x) := \frac{1}{2} x^\top P_0 x$ in $\mathbb{S}^{n-1}$, with $P_0 \in \mathbb{S}^+(n)$ constant. Since $\bar{\bar{\Gamma}}_x$ is normal, we can conclude that $H_x f(v) = \Pi_x \{ \left. \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma \right|_{t=0} \}$, with $\gamma : [-\epsilon, \epsilon] \to \mathbb{M}$ denoting a geodesic arc such that $\gamma(0) = x$ and $\dot{\gamma}(0) = v$. Since

$$\mathrm{grad}_x f = (I_n - xx^\top) P_0 x, \tag{64}$$

we have

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma = \tfrac{\mathrm{d}}{\mathrm{d}t} (P_0 \gamma - \gamma \gamma^\top P_0 \gamma) = P_0 \dot{\gamma} - \dot{\gamma} \gamma^\top P_0 \gamma - \gamma \dot{\gamma}^\top P_0 \gamma - \gamma \gamma^\top P_0 \dot{\gamma}. \tag{65}$$

Setting $t = 0$ gives

$$\left. \tfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{grad}_\gamma \right|_{t=0} = P_0 v - vx^\top P_0 x - xv^\top P_0 x - xx^\top P_0 v. \tag{66}$$

Notice that the third and the fourth addenda are normal to $T_x \mathbb{S}^{n-1}$, hence we conclude that

$$H_x f(v) = (I_n - xx^\top) P_0 v - 2vf. \tag{67}$$

In order to judge about the positive-definiteness of the Hessian, let us evaluate the inner product

$$\begin{aligned} \langle H_x f(v), v \rangle_x &= v^\top (I_n - xx^\top) P_0 v - 2vf \\ &= v^\top P_0 v - (v^\top v)(x^\top P_0 x) \\ &= \|v\|^2 (\hat{v}^\top P_0 \hat{v} - x^\top P_0 x), \end{aligned} \tag{68}$$

with $\hat{v} := v / \|v\|$ (recall that, in the positive-definiteness test, $v \neq 0$). The Hessian is positive-definite at a given point x if $\hat{v}^\top P_0 \hat{v} - x^\top P_0 x > 0$ for every $v \in T_x \mathbb{S}^2 - \{0\}$.

Recalling the expression (34), where we may set $w_\gamma := \mathrm{grad}_\gamma f$, and comparing the result with the relationship (62), it is immediate to see that they coincide, hence showing the mentioned relationship between the Hessian and the covariant derivative of a gradient field.

3.5. Axiomatization of Covariant Derivative and Relationship to Parallellism

A notion of covariant derivative may be defined arbitrarily as long as it meets generalized versions of the properties (**P$_1$**), (**P$_2$**), (**P$_3$**) recalled in Section 3.1, which qualify covariant derivative as a generalized directional derivative. Denoting the covariant derivative of a vector field $u \in \Gamma(\mathbb{M})$ in the direction $v \in T_x\mathbb{M}$ at a point $x \in \mathbb{M}$ as $\nabla_v u$, we may require it to satisfy the following properties:

$$(\mathbf{P_1}') \quad \nabla_{\alpha v + \beta w} u = \alpha \nabla_v u + \beta \nabla_w u, \tag{69}$$

$$(\mathbf{P_2}') \quad \nabla_v (u + z) = \nabla_v u + \nabla_v z, \tag{70}$$

$$(\mathbf{P_3}') \quad \nabla_v (\alpha\, u) = \alpha \nabla_v u + u\, \mathrm{d}_x \alpha(v), \tag{71}$$

with $u, z \in \Gamma(\mathbb{M})$, $v, w \in T_x\mathbb{M}$ and α, β being scalar fields and where $\mathrm{d}\alpha$ denotes a tangent map (pushforward). The properties $\mathbf{P_1}'$ and $\mathbf{P_2}'$ establish the linearity of the covariant derivative with respect to its vectorial arguments, while the property $\mathbf{P_3}'$ represents a sort of product rule.

We may verify that the connection defined in Section 3.2 satisfies the above axioms through the expression (35):

- *Proof of property* (**P$_1$'**): The covariant derivative $\nabla_{\alpha v + \beta w} u$ may be written explicitly as

$$(\nabla_{\alpha v + \beta w} u)_x = \Pi_x \{ (\partial_x w)(\alpha(x) v_x + \beta(x) w_x) + \bar{\bar{\Gamma}}_x (w_x, \alpha(x) v_x + \beta(x) w_x) \}. \tag{72}$$

  Since both the Gateaux derivative and the Christoffel form are linear in their vectorial arguments, the property follows.

- *Proof of property* (**P$_2$'**): The vector field $\nabla_v (u + z)$ may be expressed explicitly as

$$(\nabla_v (u + z))_x = \Pi_x \{ (\partial_x (u + z))(v_x) + \bar{\bar{\Gamma}}_x (u_x + z_x, v_x) \}. \tag{73}$$

  By the linearity of the terms in parentheses with respect to their vectorial arguments, the proof follows.

- *Proof of property* (**P$_3$'**): The covariant derivative $\nabla_v (\alpha\, u)$ may be re-expressed as

$$\begin{aligned}
(\nabla_v (\alpha\, u))_x &= \Pi_x \{ (\partial_x (\alpha\, u))(v_x) + \bar{\bar{\Gamma}}_x (\alpha(x) u_x, v_x) \} \\
&= \Pi_x \{ \mathrm{d}_x \alpha(v_x) u_x + \alpha(x)(\partial_x(u))(v_x) + \alpha(x) \bar{\bar{\Gamma}}_x (u_x, v_x) \}.
\end{aligned} \tag{74}$$

Noticing that $\Pi_x(u_x) \equiv u_x$ leads to the sought result.

Considering as primary the notion of covariant derivation, it is possible to recover the notion of parallelism from covariant derivation. Let us denote by $\gamma : [-\epsilon, \epsilon] \to \mathbb{M}$ a smooth curve on $\mathbb{M}$ and a smooth vector field $w \in \Gamma(\mathbb{M})$ defined (at least) on the image of γ. Such vector field is said to be *parallel along γ* if

$$\nabla_{\dot{\gamma}} w = 0. \tag{75}$$

Conversely, setting up one such differential equation equipped with an initial condition $w(0) = w_0 \in T_{\gamma(0)}\mathbb{M}$, its solution generates a *parallel transport field of the vector w_0 along the curve γ from 0 to t* denoted as $\mathbb{P}_\gamma^{0 \to t}(w_0)$.

A metric connection makes parallel transport be a *conformal isometry*. To show this important property, let us consider again a curve $\gamma : [-\epsilon, \epsilon] \to \mathbb{M}$ such that $\gamma(0) = x_0$ and two vectors $w_0, v_0 \in T_{x_0}\mathbb{M}$. Parallel transport defines two vector fields $w, v \in \Gamma(\mathbb{M})$ along such curve, namely

$$w_{\gamma(t)} := \mathbb{P}_\gamma^{0 \to t}(w_0), \quad v_{\gamma(t)} := \mathbb{P}_\gamma^{0 \to t}(v_0). \tag{76}$$

Define $s(t) := \langle w_{\gamma(t)}, v_{\gamma(t)} \rangle_{\gamma(t)}$. Since the connection ∇ is metric, the function s stays constant for any $t \in [-\epsilon, \epsilon]$. In fact, by definition of metric connection, it holds that

$$\dot{s} = \langle \nabla_{\dot{\gamma}} w_\gamma, v_\gamma \rangle_\gamma + \langle w_\gamma, \nabla_{\dot{\gamma}} v_\gamma \rangle_\gamma, \tag{77}$$

with no extra terms. Since the vector fields v_x and w_x are parallel along the curve γ, it holds that $\nabla_{\dot\gamma} w_\gamma = \nabla_{\dot\gamma} v_\gamma = 0$, which proves the assertion. A consequence of the above-proven property is that

$$\langle w_{\gamma(t)}, v_{\gamma(t)} \rangle_{\gamma(t)} = \langle w_{\gamma(0)}, v_{\gamma(0)} \rangle_{\gamma(0)} = \langle w_0, v_0 \rangle_{x_0} \tag{78}$$

for any $t \in [-\epsilon, \epsilon]$. This equality has two important consequences:

- **Parallel transport preserves the length of transported vectors**: In fact, if we take $w_0 = v_0$, we get $\langle \mathbb{P}_\gamma^{0 \to t}(v_0), \mathbb{P}_\gamma^{0 \to t}(v_0) \rangle_{\gamma(t)} = \|\mathbb{P}_\gamma^{0 \to t}(v_0)\|_{\gamma(t)}^2 = \|v_0\|_{x_0}^2$. This means that parallel transport realizes an isometry.
- **Parallel transport preserves the angle between transported vectors**: In fact, if we define the cosine of the angle between two tangent vectors as their inner product normalized by their length, we get

$$\frac{\langle \mathbb{P}_\gamma^{0 \to t}(w_0), \mathbb{P}_\gamma^{0 \to t}(v_0) \rangle_{\gamma(t)}}{\|\mathbb{P}_\gamma^{0 \to t}(w_0)\|_{\gamma(t)} \cdot \|\mathbb{P}_\gamma^{0 \to t}(v_0)\|_{\gamma(t)}} = \frac{\langle w_0, v_0 \rangle_{x_0}}{\|w_0\|_{x_0} \cdot \|v_0\|_{x_0}}, \tag{79}$$

which means that parallel transport realizes a conformal map.

The following example shows explicitly a property of parallel fields along geodesics.

Example 6. *Let us prove directly that a vector field $w \in \Gamma(\mathbb{M})$ parallel along a geodesic curve γ keeps the function $\mathcal{Q}(t) := \langle w_{\gamma(t)}, \dot\gamma(t) \rangle_{\gamma(t)}$ constant with respect to the parameter t. Recall that $\mathcal{Q} = \langle w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}}$, hence*

$$\begin{aligned} \dot{\mathcal{Q}} &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \langle w_\gamma, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \rangle^{\mathbb{A}} + \langle w_\gamma, G_\gamma(\ddot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma) \rangle^{\mathbb{A}} + \langle w_\gamma, G_\gamma(\ddot\gamma) \rangle^{\mathbb{A}}. \end{aligned} \tag{80}$$

On a geodesic line it holds that $G_\gamma(\ddot\gamma) = -\frac{1}{2} G_\gamma^\bullet(\dot\gamma, \dot\gamma) + a\ normal\ component$, hence

$$\begin{aligned} \dot{\mathcal{Q}} &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma) \rangle^{\mathbb{A}} - \frac{1}{2} \langle w_\gamma, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma) \rangle^{\mathbb{A}} - \frac{1}{2} \langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \frac{1}{2} \langle \dot\gamma, G_\gamma^\bullet(w_\gamma, \dot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \dot w_\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} + \frac{1}{2} \langle G_\gamma^{-1} G_\gamma^\bullet(w_\gamma, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \dot w_\gamma + \frac{1}{2} G_\gamma^{-1} G_\gamma^\bullet(w_\gamma, \dot\gamma), G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} \\ &= \langle \Pi_\gamma \{ \dot w_\gamma + \frac{1}{2} G_\gamma^{-1} G_\gamma^\bullet(w_\gamma, \dot\gamma) \}, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}}. \end{aligned} \tag{81}$$

Since $\bar{\bar\Gamma}_x(u, v) = \frac{1}{2} G_x^{-1} G_x^\bullet(u, v) + a\ normal\ component$, we may write that

$$\dot{\mathcal{Q}} = \langle \Pi_\gamma \{ \dot w_\gamma + \bar{\bar\Gamma}_\gamma(w_\gamma, \dot\gamma) \}, \dot\gamma \rangle_\gamma = \langle \nabla_t^\gamma w, \dot\gamma \rangle_\gamma. \tag{82}$$

Since, by assumption, the vector field w is parallel along γ, its covariant derivative $\nabla_t^\gamma w = 0$, therefore $\dot{\mathcal{Q}} = 0$, hence $\mathcal{Q}(t)$ keeps constant at its initial value $\mathcal{Q}(0)$.

3.6. Coordinate-Prone Covariant Derivation*

The fundamental requirement at the very core of a connection is that a covariant derivative of a basis vector field in a tangent direction given by another basis vector is a tangent vector itself, namely $\nabla_{\partial_i^x} \partial_j^x \in T_x\mathbb{M}$. Such a basic fact may be expressed by the fundamental relationship:

$$\nabla_{\partial_i^x} \partial_j^x = C_{ij}^k(x) \partial_k^x, \tag{83}$$

where $C_{ij}^k : \mathbb{M} \to \mathbb{R}$ are coefficients that describe the structure of a connection. The elementary covariant derivative $\nabla_{\partial_i^x} \partial_j^x$ quantifies the rate of change of the basis vector field ∂_j^x along the direction specified by the further basis vector field ∂_i^x.

By applying the properties (69)–(71), it is readily obtained that:

$$\nabla_{(v^i\partial_i^x)}(u^i\partial_i^x) = v^i\left(C_{ij}^k u^j + \frac{\partial u^k}{\partial x^i}\right)\partial_k^x, \tag{84}$$

where the functions $v^i = v^i(x)$ and the functions $u^i = u^i(x)$ are the components of vector fields in $\Gamma(\mathbb{M})$ expressed in the canonical basis $\{\partial_i^x\}$. If $C_{ij}^k \equiv 0$, the basis vectors stay constant across the tangent bundle and the covariant derivative simply coincides with the directional derivative, namely

$$C_{ij}^k \equiv 0 \;\Rightarrow\; \nabla_{(v^i\partial_i^x)}(u^i\partial_i^x) = v^i\left(\frac{\partial u^k}{\partial x^i}\right)\partial_k^x = \mathrm{d}_x u_k(v)\partial_k^x. \tag{85}$$

In general, one needs p^3 coefficient-functions to define a connection on a p-dimensional manifold, which may be specified arbitrarily and give rise to an independent definition of connection.

A special choice of connection coeffients leads to the *Levi-Civita connection*, where $C_{ij}^k := \Gamma_{ij}^k$, namely the Christoffel symbols of the second kind associated to the metric tensor of components g_{ij} which read

$$\Gamma_{ij}^k = \frac{1}{2}g^{kh}\left(\frac{\partial g_{hj}}{\partial x^i} + \frac{\partial g_{ih}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^h}\right). \tag{86}$$

Let us survey the rationale behind Levi-Civita connection through an example.

Example 7. *Let us revisit the notion of geodesic. On a manifold $\mathbb{M}$, replacing $u^k = v^k = \dot{\gamma}^k$ in the Equation (84), leads to the expression $\nabla_{\dot{\gamma}}\dot{\gamma} = \dot{\gamma}^i\dot{\gamma}^j C_{ij}^k\partial_k^x + \ddot{\gamma}^k\partial_k^x$. Setting the covariant derivative $\nabla_{\dot{\gamma}}\dot{\gamma}$ to zero leads to the geodesic equations for the components γ^k:*

$$\ddot{\gamma}^k + C_{ij}^k\dot{\gamma}^i\dot{\gamma}^j = 0. \tag{87}$$

Choosing to endow the manifold $\mathbb{M}$ with a Levi-Civita connection, namely setting $C_{ij}^k \equiv \Gamma_{ij}^k$ makes the geodesic Equation (87) coincide to the equation given in [13]. A different choice of connection would lead to a different geodesic equation and, ultimately, to a different calculus setting for the manifold $\mathbb{M}$.

3.7. Commutator of Vector Fields and Torsion Field

Let us take again two vector fields $v, w \in \Gamma(\mathbb{M})$ and a connection $\nabla : \Gamma(\mathbb{M}) \times \Gamma(\mathbb{M}) \to \Gamma(\mathbb{M})$. Apparently, in general there need not to subsist any relation between the covariant derivatives $\nabla_v w$ and $\nabla_w v$. In the case of a Levi-Civita connection, there exist however a precise relationship between such derivatives, which is usually expressed through the sentence *a Levi-Civita connection is torsion-free*.

In order to elucidate such concept, let us define two notions, namely, *commutator of two vector fields* and *torsion*. The commutator (or Lie brackets) of two vector fields $v, w \in \Gamma(\mathbb{M})$ is a further vector field denoted as

$$[v, w]_x := \Pi_x\{(\partial_x w)(v_x) - (\partial_x v)(w_x)\}, \tag{88}$$

namely, as a projected version of a comparison between the rate of change of the vector field v along the direction w_x and of the rate of change of the vector field w along the direction v_x. (As usual, the Gateaux derivatives ∂_x are taken in $\mathbb{A}$.)

The torsion field $\theta : \Gamma(\mathbb{M}) \times \Gamma(\mathbb{M}) \to \Gamma(\mathbb{M})$ is defined as

$$\theta_x(v, w) := (\nabla_v w)_x - (\nabla_w v)_x - [v, w]_x. \tag{89}$$

Torsion measures the degree of non-commutativity of directional derivations as the difference between directional derivatives of two vector fields with respect to each other depurated from their own non-commutativity. In other words, torsion measures how closely the commutator of vector fields can be recovered from the connection. Notice that, in Euclidean spaces, it holds that $(\nabla_v w)_x \equiv (\partial_x w)(v)$ and $(\nabla_w v)_x \equiv (\partial_x v)(w)$, hence torsion is certainly null. As we shall see later, null torsion induces a relationship between two otherwise unrelated concepts, namely covariant derivation and Lie derivation.

Now, let us prove that the torsion field is identically null. Let us recall that

$$(\nabla_v w)_x = \Pi_x \{\partial_x w(v_x) + \bar{\bar{\Gamma}}_x(w_x, v_x)\}, \tag{90}$$

hence

$$
\begin{aligned}
(\nabla_v w)_x - (\nabla_w v)_x &= \Pi_x \{\partial_x w(v_x)\} + \Pi_x \{\bar{\bar{\Gamma}}_x(w_x, v_x)\} \\
&\quad - \Pi_x \{\partial_x v(w_x)\} - \Pi_x \{\bar{\bar{\Gamma}}_x(v_x, w_x)\} \\
&= \Pi_x \{\partial_x w(v_x) - \partial_x v(w_x)\} \\
&= [v, w]_x,
\end{aligned}
\tag{91}
$$

since the Christoffel form is symmetric in its vectorial arguments. The assertion that a Levi-Civita connection is torsion free hence follows. We underline how the result (91) is noticeable *per se*, because it implies that, while the covariant derivative depends on the Christoffel form, the difference between two covariant derivatives with swapped arguments does not. Another way to see this property is to write $\nabla_v w = \nabla_w v + [v, w]$.

3.8. A Remarkable Relationship Linking the Operator $\Pi^\bullet$ to the Christoffel Form $\bar{\bar{\Gamma}}$

Let us recall that the bilinear operator $\Pi^\bullet : \mathbb{A} \times T\mathbb{M} \to T\mathbb{M}$ arises from the directional derivation of the orthogonal projection operator Π which, in turn, does not depend at all from the metric that a manifold $\mathbb{M}$ is endowed with, as it depends only on the structure of the tangent spaces to the manifold and on the Euclidean ambient space $\mathbb{A}$'s metric. In fact, by definition of orthogonal projector, it should hold that $\langle \Pi_x(a), v \rangle^\mathbb{A} = \langle a, v \rangle^\mathbb{A}$, for every $a \in \mathbb{A}$, $x \in \mathbb{M}$ and $v \in T_x \mathbb{M}$.

Let us recall how orthogonal projectors are computed for a few manifolds of interest.

Example 8. *Let us recall how to determine the expression of an orthogonal projector $\Pi_x : \mathbb{R}^n \to T_x \mathbb{S}^{n-1}$. By definition it must hold that $\Pi_x(a) - a \in N_x \mathbb{S}^{n-1}$, hence $\Pi_x(a) = a + \lambda x$. Since $\Pi_x(a) \in T_x \mathbb{S}^{n-1}$, it must also hold that $x^\top(a + \lambda x) = 0$, therefore $\lambda = -x^\top a$ and hence $\Pi_x(a) = a - x x^\top a$. The directional derivative of such projector along a tangent direction $v \in T_x \mathbb{S}^{n-1}$ reads*

$$\Pi_x^\bullet(a, v) = -v(x^\top a) - x(v^\top a). \tag{92}$$

Let us further recall how to derive the expression of an orthogonal projection operator $\Pi_R : \mathbb{R}^{n \times n} \to T_R \mathbb{SO}(n)$. By definition, it must hold that $\Pi_R(A) - A \in N_R \mathbb{SO}(n)$, hence $\Pi_R(A) = A + RS$, with $S = S^\top$. Since $\Pi_R(A) \in T_R \mathbb{SO}(n)$, it must also hold that $R^\top(A + RS) + (A + RS)^\top R = 0$. Straightforward calculations show that $S = -\frac{1}{2}(R^\top A + A^\top R)$ and hence $\Pi_R(A) = \frac{1}{2}(A + R A^\top R)$, where we have used the cancellation implied by $R R^\top = I_n$. The directional derivative of such projector along a tangent direction $V \in T_x \mathbb{SO}(n)$ reads

$$\Pi_R^\bullet(A, V) = -\frac{1}{2}(V A^\top R + R A^\top V). \tag{93}$$

In addition, let us survey the case of the manifold of symmetric, positive definite matrices by determining $\Pi_P : \mathbb{R}^{n \times n} \to T_P \mathbb{S}^+(n)$. By definition of orthogonal projection, it must hold that $\Pi_P(A) - A \in N_P \mathbb{S}^+(n)$, therefore $\Pi_P(A) = A + H$, with $H + H^\top = 0$. Since $\Pi_P(A) \in T_P \mathbb{S}^+(n)$, it is necessary that $A + H = A^\top - H$, therefore $H = \frac{1}{2}(A^\top - A)$, and hence $\Pi_P(A) = \frac{1}{2}(A^\top + A)$. The directional derivative of this projector along the tangent direction $V \in T_P \mathbb{S}^+(n)$ is identically zero, namely

$$\Pi_P^\bullet(A, V) \equiv 0, \tag{94}$$

since $\Pi_P(A)$ does not actually depend on the foot-point P.

To end this example with, let us consider the case of the Stiefel manifold $\mathrm{St}(n, p)$. In this case, $\Pi_X : \mathbb{R}^{n \times p} \to T_X\mathrm{St}(n, p)$. From the condition $\Pi_X(A) - A \in N_X\mathrm{St}(n, p)$ it follows that $\Pi_X(A) = A + XS$, with $S = S^\top$. From the tangency condition $\Pi_X(A) \in T_X\mathrm{St}(n, p)$ it follows that $X^\top(A + XS) + (A + XS)^\top X = 0$, hence that $S = -\frac{1}{2}(X^\top A + A^\top X)$. Ultimately, we get $\Pi_X(A) = A - \frac{1}{2}X(X^\top A + A^\top X)$. The directional derivative of such projector along a tangent direction $V \in T_X\mathrm{St}(n, p)$ reads

$$\Pi_X^\bullet(A, V) = -\tfrac{1}{2}X(V^\top A + A^\top V) - \tfrac{1}{2}V(X^\top A + A^\top X), \tag{95}$$

that the expression (93) is a special case of, when $p = n$.

We further recall that the Christoffel form of the second kind $\bar{\bar{\Gamma}} : (T\mathbb{M})^2 \to \mathbb{A}$ does depend on the metric, through the metric kernel $G : T\mathbb{M} \to T\mathbb{M}$, and from the structure of the tangent bundle $T\mathbb{M}$. Let us recall the expressions of some Christoffel forms of interest.

Example 9. *For the hyper-sphere $\mathbb{S}^{n-1}$ endowed with the canonical metric, it holds*

$$\bar{\bar{\Gamma}}_x(u, v) = x(u^\top v). \tag{96}$$

(We note, with apologies, that in the first part of this series of tutorial [13], the Christoffel form associated to the unit hyper-sphere was reported with an incorrect sign).

For the special orthogonal group, it holds

$$\bar{\bar{\Gamma}}_R(U, V) = \tfrac{1}{2}R(U^\top V + V^\top U). \tag{97}$$

For the manifold of symmetric, positive definite matrices, the Christoffel form takes the expression

$$\bar{\bar{\Gamma}}_P(U, V) = \tfrac{1}{2}UP^{-1}V - \tfrac{1}{2}VP^{-1}U. \tag{98}$$

Instead, the expression

$$\bar{\bar{\Gamma}}_X(U, V) = \tfrac{1}{2}X(U^\top V + V^\top U) \tag{99}$$

pertains to the Stiefel manifold endowed with the Euclidean metric.

Now we lay out a quite simple observation, that leads to a number of important consequences. Let us consider a smooth vector field $w \in \Gamma(\mathbb{M})$, namely, a function that assigns a tangent vector $w_x \in T_x\mathbb{M}$ to every point $x \in \mathbb{M}$, and a smooth curve $\rho : [-\epsilon, \epsilon] \to \mathbb{M}$. In the previous sections we had familiarized ourselves with the naïve directional derivative $\dot{w}_\rho$, that actually stands for $\frac{\mathrm{d}}{\mathrm{d}t}w_{\rho(t)}$. Now the fundamental observation that we are underlining is that, although $\dot{w}_\rho$ is an array in the ambient space $\mathbb{A}$, *this does not mean that its entries are unconstrained*. In other terms, the naïve directional derivative of a tangent vector field is not a tangent vector field, yet it must obey some constraints.

Let us express such constraints in an analytic way. For the sake of notation conciseness, we are going to make use of the following notation: For any $a \in \mathbb{A}$, we shall denote by $a_x^\perp$ the quantity $a - \Pi_x(a)$.

Let us observe that, by definition of tangent vector field, on a smooth curve $\rho : [-\epsilon, \epsilon] \to \mathbb{M}$ it must hold that $\Pi_\rho(w_\rho) = w_\rho$. Deriving with respect to the parameter t gives:

$$\Pi_\rho^\bullet(w_\rho, \dot{\rho}) + \Pi_\rho(\dot{w}_\rho) = \dot{w}_\rho. \tag{100}$$

Setting $t = 0$, $x := \rho(0)$ and $v := \dot{\rho}(0)$, gives

$$[\partial_x w(v)]^\perp = \Pi_x^\bullet(w_x, v). \tag{101}$$

The above result means that the normal component of the naïve velocity of a vector field is completely determined by the directional derivative $\Pi^\bullet$ (hence only its tangential component may be prescribed freely). (This finding further clarifies the meaning of the bilinear operator $\Pi^{\bullet}$.) It is worth underlining that the relation (101) only involves the restriction $\Pi^\bullet$ in $(T\mathbb{M})^2$ and is fully compliant with the previous finding that $\Pi^\bullet : (T\mathbb{M})^2 \to N\mathbb{M}$.

An important consequence of the above finding is as follows. We know that a tangent vector field $w \in \Gamma(\mathbb{M})$ is parallel along a curve $\rho : [-\epsilon, \epsilon] \to \mathbb{M}$ if it satisfies the differential equation (in $\mathbb{A}$)

$$\dot{w}_\rho + \bar{\bar{\Gamma}}_\rho(w_\rho, \dot{\rho}) = 0. \tag{102}$$

Splitting the members of the above equation into their tangential and normal components, and retaining the latter, we get, for a point $x := \rho(0)$:

$$[\partial_x w(v)]^\perp = -\bar{\bar{\Gamma}}_x^\perp(w_x, v). \tag{103}$$

where $v := \dot{\rho}(0)$. Comparing the relationship (101) with the relation (103), we immediately get that

$$\Pi_x^\bullet(u, v) + \bar{\bar{\Gamma}}_x^\perp(u, v) = 0, \text{ for every } u, v \in T_x\mathbb{M}, \tag{104}$$

namely, the anti-normal component of the Christoffel form (of the second kind) equals a restriction of the directional derivative of the projection operator. Two seemingly unrelated operators are quite tightly bound to one another!

Let us verify the above important finding on a number of cases.

Example 10. *For the hyper-sphere $\mathbb{S}^{n-1}$ endowed with the canonical metric, the restriction of the directional derivative $\Pi^\bullet$ to the tangent bundle reads*

$$\Pi_x^\bullet(u, v) = -v(x^\top u) - x(v^\top u) = -x(u^\top v), \tag{105}$$

because $x^\top u = 0$. Notice that, in this example, $\bar{\bar{\Gamma}}^\perp \equiv \bar{\bar{\Gamma}}$, hence it coincides with the Christoffel form (96) up to a sign switch.

For the special orthogonal group endowed with its canonical metric, the restriction of the directional derivative $\Pi^\bullet$ to the tangent bundle reads

$$\begin{aligned} \Pi_R^\bullet(U, V) &= -\tfrac{1}{2}VU^\top R - \tfrac{1}{2}RU^\top V \\ &= +\tfrac{1}{2}VR^\top U - \tfrac{1}{2}RU^\top V \\ &= -\tfrac{1}{2}RV^\top U - \tfrac{1}{2}RU^\top V. \end{aligned} \tag{106}$$

To conduct the above calculations, we have used twice the property that $V^\top R = -R^\top V$ for every $V \in T_R\mathbb{SO}(n)$. Notice that, even in this case, $\bar{\bar{\Gamma}}^\perp \equiv \bar{\bar{\Gamma}}$, hence $\Pi_R^\bullet(U, V)$ coincides, up to a sign, with the Christoffel form (97).

In the case of the manifold of symmetric, positive definite matrices endowed with its canonical metric, the directional derivative of the orthogonal projector, although restricted to the tangent bundle, is still identically null, namely $\Pi_P^\bullet(U, V) \equiv 0$. On the other hand, from the expression (98), it is immediately verified that $\bar{\bar{\Gamma}}^\perp \equiv 0$, hence the two expressions coincide.

For the Stiefel manifold, the directional derivative of the orthogonal projector restricted to the tangent bundle reads

$$\begin{aligned} \Pi_X^\bullet(U, V) &= -\tfrac{1}{2}X(V^\top U + U^\top V) - \tfrac{1}{2}V(X^\top U + U^\top X) \\ &= -\tfrac{1}{2}X(V^\top U + U^\top V), \end{aligned} \tag{107}$$

since the second term on the right-hand side cancels. Once again, this quantity, which is purely normal, coincides (up to sign) with the Christoffel form of the second kind (99) which, in turn, appears to be purely normal.

Notice that in all these examples, the expression of $\Pi^\bullet$ results to be symmetric in its vectorial arguments.

The property (104) prescribes specifically the normal component of the Christoffel form of the second kind to be independent of the metric. Such prescription implies several important consequences, that we are going to explore in the following.

A consequence to underline is that $\Pi_x^\bullet(u, v) = \Pi_x^\bullet(v, u)$ for every pair $u, v \in T_x\mathbb{M}$, namely, the restriction of the directional derivative of the orthogonal projector to the tangent bundle is symmetrical in its arguments.

A further consequence is that, although for a non-parallel vector field $w \in \Gamma(\mathbb{M})$ along a curve ρ it holds that $\dot{w}_\rho + \bar{\bar{\Gamma}}_\rho(w_\rho, \dot{\rho}) \neq 0$, from the relationships (101) and (104) it follows that

$$\dot{w}_\rho^\perp + \bar{\bar{\Gamma}}_\rho^\perp(w_\rho, \dot{\rho}) \equiv 0 \text{ for } every \text{ vector field } w \in \Gamma(\mathbb{M}) \text{ along } \rho. \tag{108}$$

This latter relationship also implies that the normal component of the naïve velocity of variation of a (non-parallel) vector field may be written as a function of the vector field itself.

As a further consequence, let us examine the expression of the commutator of two vector fields as defined in Equation (88). Given two vector fields $u, v \in \Gamma(\mathbb{M})$, their commutator at a point x is given by

$$[u, v]_x = \Pi_x\{\partial_x v(u_x)\} - \Pi_x\{\partial_x u(v_x)\}. \tag{109}$$

Since for the two vector fields in must hold that

$$\begin{cases} \Pi_x(\partial_x v(u_x)) = \partial_x v(u_x) - \Pi_x^\bullet(v_x, u_x), \\ \Pi_x(\partial_x u(v_x)) = \partial_x u(v_x) - \Pi_x^\bullet(u_x, v_x), \end{cases} \tag{110}$$

by the symmetry of the operator $\Pi^\bullet$ restricted to the tangent bundle it follows that

$$[u, v]_x = \partial_x v(u_x) - \partial_x u(v_x). \tag{111}$$

An even deeper consequence of the property (104) concerns the computation of the covariant derivative of a vector field. Let us recall that, given a vector field $w \in \Gamma(\mathbb{M})$ and a tangent vector $v \in T_x\mathbb{M}$, the covariant derivative of w in the direction v at a point x is given by

$$\begin{aligned} (\nabla_v w)_x &= \Pi_x\{\partial_x w(v)\} + \Pi_x\{\bar{\bar{\Gamma}}_x(w_x, v)\} \\ &= \partial_x w(v) - \Pi_x^\bullet(w_x, v) + \bar{\bar{\Gamma}}_x(w_x, v) - \bar{\bar{\Gamma}}_x^\perp(w_x, v). \end{aligned} \tag{112}$$

Since the second and the fourth terms in the above sum cancel out, we rest with

$$(\nabla_v w)_x = \partial_x w(v) + \bar{\bar{\Gamma}}_x(w_x, v). \tag{113}$$

In other terms, the sum $\partial_x w(v) + \bar{\bar{\Gamma}}_x(w_x, v)$ is already an element of the tangent bundle, without any need for projection. A similar relationship holds for the covariant derivative along a smooth curve, namely

$$\nabla_t^\gamma w = \partial_\gamma w(\dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}). \tag{114}$$

The above two expressions of the covariant derivative may be used interchangeably with those given in Section 3.2 whenever they are deemed to be easier to apply, as in the case of iterated derivatives covered in Section 4.2.

Let us verify the above general relations by a number of examples.

Example 11. *Let us consider, as a first case of study, the computation of the covariant derivative on the unit hyper-sphere $\mathbb{S}^{n-1}$ endowed with the canonical metric. In Example 3 we have evaluated the covariant derivative $\nabla_t^\gamma w$ of a vector field $w \in \Gamma(S^{n-1})$ in two different ways to prove their equivalency. Let us repeat such calculation by means of the relationship (114):*

$$\nabla_t^\gamma w = \dot{w}_\gamma + (w_\gamma^\top \dot{\gamma})\gamma. \tag{115}$$

Now, since $\gamma^\top w_\gamma = 0$, deriving once gives $\dot{\gamma}^\top w_\gamma + \gamma^\top \dot{w}_\gamma = 0$, therefore

$$\nabla_t^\gamma w = \dot{w}_\gamma - (\gamma^\top \dot{w}_\gamma)\gamma = (I_n - \gamma\gamma^\top)\dot{w}_\gamma, \tag{116}$$

as in Example 3.

The case regarding the manifold of symmetric, positive-definite matrices endowed with the canonical metric is self-explanatory, since the Christoffel form of the second kind is purely tangential.

The case concerning the special orthogonal group is more interesting. In particular, in [17] (Theorem 1), the present author gave the following expression for the covariant derivative of a vector field $W \in \Gamma(\mathbb{SO}(n))$ along a curve $R(t)$:

$$\nabla_t^R W = (\dot{R}R^\top + \tfrac{3}{2}R\dot{R}^\top)W_R + \dot{W}_R + \tfrac{1}{2}W_R\dot{R}^\top R. \tag{117}$$

Such expression was derived applying directly the definition in terms of parallel transport. We would like to show here that such expression is exactly the same as that found in Example 3, but applying the Equation (114). Let us first simplify the expression (117) for the sake of easier comparison. From the orthogonality condition $RR^\top = I_n$ it follows that $\dot{R}R^\top + R\dot{R}^\top = 0$, hence $\dot{R}R^\top + \tfrac{3}{2}R\dot{R} = \tfrac{1}{2}R\dot{R}^\top$, therefore the expression (117) simplifies in

$$\nabla_t^R W = \tfrac{1}{2}R\dot{R}^\top W_R + \dot{W}_R + \tfrac{1}{2}W_R\dot{R}^\top R. \tag{118}$$

Further we notice that from the condition $RW_R^\top + W_RR^\top = 0$ it follows that $W_R\dot{R}^\top = -\dot{R}W_R^\top - R\dot{W}_R^\top - \dot{W}_RR^\top$, hence

$$\begin{aligned}
\nabla_t^R W &= \tfrac{1}{2}R\dot{R}^\top W_R + \dot{W}_R - \tfrac{1}{2}(\dot{R}W_R^\top + R\dot{W}_R^\top + \dot{W}_RR^\top)R \\
&= \tfrac{1}{2}R\dot{R}^\top W_R + \dot{W}_R - \tfrac{1}{2}\dot{R}W_R^\top R - \tfrac{1}{2}R\dot{W}_R^\top R - \tfrac{1}{2}\dot{W}_R \\
&= \tfrac{1}{2}\dot{W}_R - \tfrac{1}{2}R\dot{W}_R^\top R,
\end{aligned} \tag{119}$$

in fact $R\dot{R}^\top W_R = -\dot{R}R^\top W_R = \dot{R}W_R^\top R$, hence the first and the third terms in the second line cancel out. This is the expression found in Example 3. By applying the simplified Equation (114) we get:

$$\nabla_t^R W = \dot{W}_R + \tfrac{1}{2}RW_R^\top \dot{R} + \tfrac{1}{2}R\dot{R}^\top W_R. \tag{120}$$

On the other hand, from the tangency condition $W_R^\top R + R^\top W_R = 0$ it follows that $W_R^\top \dot{R} = -\dot{W}_R^\top R - \dot{R}^\top W_R - R^\top \dot{W}_R$, hence

$$\nabla_t^R W = \dot{W}_R - \tfrac{1}{2}R\dot{W}_R^\top R - \tfrac{1}{2}R\dot{R}^\top W_R - \tfrac{1}{2}RR^\top \dot{W}_R + \tfrac{1}{2}R\dot{R}^\top W_R \tag{121}$$

which, upon easy simplifications, come to coincide with expression (119).

The projectionless expression of the covariant derivative of a vector field along a curve impacts on a number of differential operators, including the Hessian, which is defined as the covariant derivative of a gradient of a function. In fact, let us recall that, given a smooth function $f : \mathbb{M} \to \mathbb{R}$, its Hessian operator computed along a curve γ is defined as $H_\gamma f(\dot{\gamma}) := \nabla_t^\gamma \text{grad}_\gamma f$, where now

$$\begin{cases} \nabla_t^\gamma \text{grad}_\gamma f = \frac{\mathrm{d}}{\mathrm{d}t}\text{grad}_\gamma f + \bar{\bar{\Gamma}}_\gamma(\text{grad}_\gamma f, \dot{\gamma}), \\ \text{grad}_\gamma f = G_\gamma^{-1}(\Pi_\gamma\{\partial_\gamma f\}), \end{cases} \tag{122}$$

hence an alternative expression for the Hessian is

$$H_\gamma f(\dot\gamma) = \tfrac{\mathrm{d}}{\mathrm{d}t} G_\gamma^{-1}(\Pi_\gamma\{\partial_\gamma f\}) + \bar{\bar\Gamma}_\gamma(G_\gamma^{-1}(\Pi_\gamma\{\partial_\gamma f\}), \dot\gamma). \tag{123}$$

The following examples studies in particular two interesting problems involving covariant differentiation and the logarithmic map.

Example 12. *As a first example of relationship involving covariant differentiation and logarithmic map, let us consider the covariant derivative of a logarithmic map field. In particular, given a point $x \in \mathbb{M}$ and a smooth curve γ, the quantity $\log_\gamma x$ represents a vector field along the curve γ. One might wonder how the covariant derivative of such vector field along the curve γ looks like. Applying the Equation (114) to the vector field $w_\gamma := \log_\gamma x$ (where x is fixed), we get*

$$\nabla_t^\gamma \log_\gamma x = (\partial_\gamma \log_\gamma x)(\dot\gamma) + \bar{\bar\Gamma}_\gamma(\log_\gamma x, \dot\gamma). \tag{124}$$

Technically speaking, one might notice that $\partial_y \log_y x \in \mathrm{End}(T_y\mathbb{M})$ and that it might be termed as logarithmic pushforward endomorphism. *Such endomorphism was mentioned also in [18].*

For the second example, let us consider a manifold $\mathbb{M}_2$ of dimension 2 endowed with a distance function $\mathscr{D} : \mathbb{M}_2 \times \mathbb{M}_2 \to \mathbb{R}$, take a point $x \in \mathbb{M}_2$ and define a curve γ implicitly such as $\mathscr{D}^2(\gamma, x) = r^2$ for some $r > 0$. Notice that γ is a curve only if the manifold is of dimension two (otherwise it would be a 'surface' or, more generally, a submanifold) and that such curve is clearly a 'circle', centered in x, of radius r. In addition, let us choose a special parametrization of the curve γ known as parametrization by arclength, *which essentially means that at any point of the curve it presents unit speed, namely $\|\dot\gamma\|_\gamma = 1$. Now one might wonder as to whether the acceleration of the curve $\nabla_t^\gamma \dot\gamma$ presents any special property (where ∇ denotes the Levi-Civita connection). This is indeed the case: let us examine the problem. As a first consideration, notice that the condition of unit speed implies that $\nabla_t^\gamma \dot\gamma \perp \dot\gamma$, in fact*

$$\langle \dot\gamma, \dot\gamma \rangle_\gamma = \mathrm{constant} \overset{\mathrm{d}/\mathrm{d}t}{\Rightarrow} 2\langle \nabla_t^\gamma \dot\gamma, \dot\gamma \rangle_\gamma = 0 \Rightarrow \nabla_t^\gamma \dot\gamma \perp \dot\gamma \tag{125}$$

on each tangent space $T_\gamma\mathbb{M}$. In addition, the condition $\mathscr{D}^2(\gamma, x) = r^2$ implies that

$$0 = \tfrac{\mathrm{d}}{\mathrm{d}t} \mathscr{D}^2(\gamma, x) = \langle -2\log_\gamma x, \dot\gamma \rangle_\gamma \Rightarrow \log_\gamma x \perp \dot\gamma. \tag{126}$$

Now, since at every point of the curve $\dim(T_\gamma M) = 2$, one must conclude that

$$\nabla_t^\gamma \dot\gamma \parallel \log_\gamma x, \tag{127}$$

namely the covariant acceleration is directed radially. As a by-product, with reference to the first example, we may notice that the condition $\mathscr{D}^2(\gamma, x) = r^2$ also implies that

$$0 = \tfrac{\mathrm{d}}{\mathrm{d}t}\langle \log_\gamma x, \log_\gamma x \rangle_\gamma = 2\langle \nabla_t^\gamma \log_\gamma x, \log_\gamma x \rangle_\gamma \Rightarrow \nabla_t^\gamma \log_\gamma x \perp \log_\gamma x, \tag{128}$$

which therefore implies

$$\nabla_t^\gamma \log_\gamma x \parallel \dot\gamma. \tag{129}$$

In fact we may notice that, for a Euclidean space, it holds that $\log_\gamma x \equiv x - \gamma$ and hence it follows that $\nabla_t^\gamma \log_\gamma x = -\dot\gamma$.

To end this section, let us examine a further, perhaps counter-intuitive, example of calculation.

Example 13. *We have seen that an important property of the Levi-Civita connection is that is is 'metric', which may be expressed concisely as $\tfrac{\mathrm{d}}{\mathrm{d}t}\langle \star, \star \rangle_{\gamma(t)} \equiv 0$ for any smooth curve γ. In intrinsic manifold calculus, such property is usually expressed through the sentence 'the covariant derivative*

of the metric tensor is equal to zero'. One might thus wonder as to whether the same holds for the metric kernel G, namely if $\nabla_t^\gamma G_\gamma \overset{?}{\equiv} 0$. We are going to show now that this is not, indeed, the case.

Recalling that, for a smooth vector field $w \in \Gamma(\mathbb{M})$ it holds that $\nabla_t^\gamma w = \dot{w}_\gamma + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma})$, we may compute

$$
\begin{aligned}
\nabla_t^\gamma G_\gamma(w_\gamma) &= \tfrac{\mathrm{d}}{\mathrm{d}t} G_\gamma(w_\gamma) + \bar{\bar{\Gamma}}_\gamma(G_\gamma(w_\gamma), \dot{\gamma}) \\
&= G_\gamma^\bullet(w_\gamma, \dot{\gamma}) + G_\gamma(\dot{w}_\gamma) + \bar{\bar{\Gamma}}_\gamma(G_\gamma(w_\gamma), \dot{\gamma}) \\
&= G_\gamma(\nabla_t^\gamma w) + G_\gamma^\bullet(w_\gamma, \dot{\gamma}) - G_\gamma(\bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma})) + \bar{\bar{\Gamma}}_\gamma(G_\gamma(w_\gamma), \dot{\gamma}).
\end{aligned}
\tag{130}
$$

Therefore, we can conclude that the covariant derivative of the metric kernel reads $\nabla_t^\gamma G_\gamma = G_\gamma^\bullet(\star, \dot{\gamma}) - G_\gamma(\bar{\bar{\Gamma}}_\gamma(\star, \dot{\gamma})) + \bar{\bar{\Gamma}}_\gamma(G_\gamma(\star), \dot{\gamma})$, where $\star$ denotes an empty slot again. The expression on the right-hand side of the expression (130) is equal to zero trivially only when the metric kernel coincides with the identity.

3.9. Lie Derivation*

Given two vector fields $u, v \in \Gamma(\mathbb{M})$, besides of covariant derivation $\nabla_u v$, there exists a further way to compute the rate of change of v in the direction u, which constitutes the first attempt to define directional derivatives on manifolds: the Lie derivative. Lie derivation does not require any notion of connection, parallel transport nor metrization. We shall however see that Lie derivation and Levi-Civita covariant derivation are related to one another. Although the notion of Lie derivative is presented in this section in extrinsic terms, it is a non-essential topic in this tutorial paper. However, besides of its historical value, we would like to point out that it has important applications in control theory, as summarized in [19].

Let $u, v \in \Gamma(\mathbb{M})$ denote two smooth tangent vector fields. The notation u_x indicates the value of u at x. The Lie derivative of a vector field is defined in an analogous way to the covariant derivative, except that the role of parallel transport is played by a pullback map of a flow. We define a flow $\sigma_t(x)$ associated to the vector field u as the solution of the following differential equation:

$$
\dot{\sigma}_t(x) = u_{\sigma_t(x)}, \text{ with initial condition } \sigma_0(x) = x.
\tag{131}
$$

The function $\sigma_t : \mathbb{M} \to \mathbb{M}$ takes a point in the manifold to another point in the manifold, hence its first-order approximation at a point x, given by its tangent (pushforward) map, may be defined as

$$
\mathrm{d}_x \sigma_t(v) := \left. \frac{\mathrm{d}}{\mathrm{d}s} \sigma_t(\tau_s(x)) \right|_{s=0},
\tag{132}
$$

where $\tau : [-\epsilon, \epsilon] \to \mathbb{M}$ denotes any smooth curve such that $\tau_0(x) = x$ and $\dot{\tau}_0(x) = v_x \in T_x\mathbb{M}$. Now, the pushforward map $\mathrm{d}_x \sigma_t : T_x\mathbb{M} \to T_{\sigma_t(x)}\mathbb{M}$, hence the associated pullback $(\mathrm{d}_x \sigma_t)^{-1} : T_{\sigma_t(x)}\mathbb{M} \to T_x\mathbb{M}$. Namely, the pullback $(\mathrm{d}_x \sigma_t)^{-1}$ takes a vector from the tangent space $T_{\sigma_t(x)}\mathbb{M}$ to the tangent space $T_x\mathbb{M}$. Such operation resembles parallel transport but is however referred to as *Lie dragging*. Lie derivation is based on Lie dragging in place of parallel transport and is defined as

$$
(\mathcal{L}_u v)_x := \left. \frac{\mathrm{d}}{\mathrm{d}t} (\mathrm{d}_x \sigma_t)^{-1} v_{\sigma_t(x)} \right|_{t=0}.
\tag{133}
$$

We notice that the Lie derivative is also related to the differential of a scalar-valued function $f : \mathbb{M} \to \mathbb{R}$, since $\mathrm{d}_x f(v) = (\mathcal{L}_v f)_x$.

In order to find a rule to calculate the Lie derivative of a vector field, we can proceed as follows. By definition we have

$$
\begin{aligned}
(\mathcal{L}_u v)_x &= \lim_{h \to 0} \frac{(\mathrm{d}_x \sigma_h)^{-1} v_{\sigma_h(x)} - v_x}{h} \\
&= \lim_{h \to 0} (\mathrm{d}_x \sigma_h)^{-1} \frac{v_{\sigma_h(x)} - (\mathrm{d}_x \sigma_h) v_x}{h} \\
&= \lim_{h \to 0} \frac{v_{\sigma_h(x)} - (\mathrm{d}_x \sigma_h) v_x}{h} \\
&= \lim_{h \to 0} \frac{v_{\sigma_h(x)} - \frac{\mathrm{d}}{\mathrm{d}s} \sigma_h(\tau_s) \big|_{s=0}}{h},
\end{aligned}
\tag{134}
$$

where we used the fact that $(\mathrm{d}_x \sigma_0)^{-1} = \mathrm{id}_x$ and the last equality is due to the definition (132). By Taylor-expanding the function $\sigma_t(x)$ in t we get

$$
\sigma_t(x) = x + t \frac{\mathrm{d}}{\mathrm{d}t} \sigma_t(x) \Big|_{t=0} + \mathcal{O}(t^2) = x + t\, u_x + \mathcal{O}(t^2).
\tag{135}
$$

Hence, it follows that

$$
\sigma_t(\tau_s(x)) = \tau_s(x) + t\, u_{\tau_s(x)} + \mathcal{O}(t^2)
\tag{136}
$$

and, consequently, that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} \sigma_t(\tau_s(x)) &= \dot{\tau}_s(x) + t \frac{\mathrm{d}}{\mathrm{d}s} u_{\tau_s(x)} + \mathcal{O}(t^2) \\
&= \dot{\tau}_s(x) + t\, (\partial_x u)(\dot{\tau}_s(x)) + \mathcal{O}(t^2).
\end{aligned}
\tag{137}
$$

Setting $s = 0$ results in

$$
\frac{\mathrm{d}}{\mathrm{d}s} \sigma_t(\tau_s(x)) \Big|_{s=0} = \dot{\tau}_0(x) + t\, (\partial_x u)(\dot{\tau}_0(x)) + \mathcal{O}(t^2) = v_x + t\, (\partial_x u)(v_x) + \mathcal{O}(t^2).
\tag{138}
$$

Likewise, Taylor-expanding the function $v_{\sigma_t(x)}$ with respect to the parameter t gives

$$
v_{\sigma_t(x)} = v_x + t\, (\partial_x v)(\dot{\sigma}_0(x)) + \mathcal{O}(t^2) = v_x + t\, (\partial_x v)(u_x) + \mathcal{O}(t^2).
\tag{139}
$$

Gluing pieces together gives

$$
(\mathcal{L}_u v)_x = \partial_x v(u_x) - \partial_x u(v_x) = [u, v]_x.
\tag{140}
$$

Namely, the Lie derivative of a vector field along a given direction coincides with a commutator.

Next we discuss, as an example, the meaning of vanishing Lie derivative.

Example 14. *Given two vector fields $u, v \in \Gamma(\mathbb{M})$, it may happen that $\mathcal{L}_u v = 0$. Intuitively, such condition means that the two vector fields are related to one another. In particular, recalling the meaning of the flow σ at the basis of the definition (133), mathematicians say that σ is a 'symmetry transformation' of the vector field v. In order to make such notion explicit, let us rewrite the expression (133) for a generic point of the flow, namely*

$$
(\mathcal{L}_u v)_{\sigma_t(x)} := \frac{\mathrm{d}}{\mathrm{d}t} (\mathrm{d}_x \sigma_t)^{-1} v_{\sigma_t(x)}.
\tag{141}
$$

If the Lie derivative of the vector field v vanishes along the whole flow of the vector field u, the above expression implies that the function $t \mapsto (\mathrm{d}_x \sigma_t)^{-1} v_{\sigma_t(x)}$ is constant, hence that $(\mathrm{d}_x \sigma_t)^{-1} v_{\sigma_t(x)} = (\mathrm{d}_x \sigma_0)^{-1} v_{\sigma_0(x)}$. Ultimately, a vanishing Lie derivatives implies that

$$
v_{\sigma_t(x)} = (\mathrm{d}_x \sigma_t) v_x,
\tag{142}
$$

namely that the vector field v along the flow takes a value obtainable by Lie-dragging backward its source value and is completely 'generated' by it. This is reminiscent of the construction of a parallel vector field, from an initial seed, by parallel transport.

A consequence of the above result is that a torsion-free connection is nicely related to the Lie derivative, in fact, the condition $\theta(u, v) = 0$ implies that

$$\nabla_u v - \nabla_v u = \mathcal{L}_u v \tag{143}$$

while, in general, connections need not be related to Lie derivatives. Notice that, from the above expression, it directly follows that $\mathcal{L}_u v + \mathcal{L}_v u = 0$. But there is more to this. In fact, upon introducing a metric and a notion of Lie derivation, a covariant derivative may be computed on the basis of Lie derivatives through the following formula due to Koszul:

$$2\langle \nabla_v u, w \rangle = \mathcal{L}_v \langle u, w \rangle + \mathcal{L}_u \langle v, w \rangle - \mathcal{L}_w \langle v, u \rangle - \langle u, \mathcal{L}_v w \rangle - \langle v, \mathcal{L}_u w \rangle - \langle w, \mathcal{L}_u v \rangle, \tag{144}$$

where $u, v, w \in \Gamma(\mathbb{M})$ are smooth vector fields and we have omitted for the ease of notation the point $x \in \mathbb{M}$ where all derivatives are taken. Proving the above identity is a matter of patience. First, let us recall that given a smooth curve $\gamma : [-\epsilon, \epsilon] \to \mathbb{M}$ such that $\gamma(0) = x \in \mathbb{M}$ and $\dot\gamma(0) = v_x \in T_x\mathbb{M}$, and given two smooth vector fields $u, w \in \Gamma(\mathbb{M})$, for a metric connection it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle u_{\gamma(t)}, w_{\gamma(t)} \rangle_{\gamma(t)} = \langle \nabla_t^\gamma u, w_{\gamma(t)} \rangle_{\gamma(t)} + \langle u_{\gamma(t)}, \nabla_t^\gamma w \rangle_{\gamma(t)}. \tag{145}$$

Setting $t = 0$ leads to

$$(\mathrm{d}_x \langle u_x, w_x \rangle_x)(v_x) = (\mathcal{L}_v \langle u, w \rangle)_x = \langle (\nabla_v u)_x, w_x \rangle_x + \langle u_x, (\nabla_v w)_x \rangle_x. \tag{146}$$

Likewise, we may write that

$$\begin{aligned}
(\mathcal{L}_u \langle v, w \rangle)_x &= \langle (\nabla_u v)_x, w_x \rangle_x + \langle v_x, (\nabla_u w)_x \rangle_x, \\
-(\mathcal{L}_w \langle v, u \rangle)_x &= -\langle (\nabla_w v)_x, u_x \rangle_x - \langle v_x, (\nabla_w u)_x \rangle_x.
\end{aligned} \tag{147}$$

Summing up side by side yields:

$$\begin{aligned}
&(\mathcal{L}_v \langle u, w \rangle)_x + (\mathcal{L}_u \langle v, w \rangle)_x - (\mathcal{L}_w \langle v, u \rangle)_x \\
&= \langle (\nabla_v u)_x + (\nabla_u v)_x, w_x \rangle_x + \langle (\nabla_v w)_x - (\nabla_w v)_x, u_x \rangle_x \\
&\quad + \langle (\nabla_u w)_x - (\nabla_w u)_x, v_x \rangle_x \\
&= \langle (\nabla_v u)_x + (\nabla_v u)_x + (\mathcal{L}_u v)_x, w_x \rangle_x + \langle (\mathcal{L}_v w)_x, u_x \rangle_x + \langle (\mathcal{L}_u w)_x, v_x \rangle_x \\
&= 2\langle (\nabla_v u)_x, w_x \rangle_x + \langle (\mathcal{L}_u v)_x, w_x \rangle_x + \langle (\mathcal{L}_v w)_x, u_x \rangle_x + \langle (\mathcal{L}_u w)_x, v_x \rangle_x,
\end{aligned} \tag{148}$$

which proves the assertion.

To conclude this stub, we may examine an example that discusses the notion of Killing vector field.

Example 15. *A* Killing vector field *is one along with the Lie derivative of the metric vanishes. Let us make this notion more substantial. Let $u, v, w \in \Gamma(\mathbb{M})$ denote three vector fields and let us recall that $\mathcal{L}_v \langle u, w \rangle = \langle \nabla_v u, w \rangle + \langle u, \nabla_v w \rangle$, where ∇ denotes a Levi-Civita connection. Now, from the relationship (143) it follows that*

$$\nabla_v u = \nabla_u v + \mathcal{L}_v u, \ \nabla_v w = \nabla_w v + \mathcal{L}_v w, \tag{149}$$

therefore, it holds that

$$\underbrace{\mathcal{L}_v\langle u,w\rangle - \langle\mathcal{L}_v u,w\rangle - \langle u,\mathcal{L}_v w\rangle}_{:=\mathcal{L}_v\langle\star,\star\rangle} = \langle\nabla_u v,w\rangle + \langle u,\nabla_w v\rangle. \tag{150}$$

The left-hand side of such equation represents the Lie derivative of the metric, namely $\mathcal{L}_v\langle\star,\star\rangle$. Vanishing of the Lie derivative of the metric implies that

$$\langle\nabla_\star v,\star\rangle + \langle\star,\nabla_\star v\rangle = 0, \tag{151}$$

that is a condition on the vector field v referred to as Killing equation. *Here, the empty slots remind us that the equation $\langle\nabla_u v,w\rangle + \langle u,\nabla_w v\rangle = 0$ must hold for every choice of $u,w \in \Gamma(\mathbb{M})$, hence the Killing equation involves only the differential structure of the manifold.*

4. Iterated Derivatives and Riemannian Curvature

The notion of covariant derivative arises as a generalization of the directional derivative of a vector field. However, while directional derivation commutes, covariant derivation does not. Such observation leads to the notion of Riemannian curvature endomorphism.

The present section is organized as follows. Section 4.1 remarks some well-known properties of the iterated directional derivatives in the flat space $\mathbb{R}^n$. The Section 4.2 lays out the notion of iterated covariant derivatives and defines the notion of second covariant derivative, which is shown to be well-defined in contrast to a double iterated derivative. Section 4.3 describes the parallel transport of a tangent vector along a tiny loop and shows that the final result differs from the starting vector, and attributes such discrepancy to the curvedness of a manifold. In Section 4.4, we describe the key notion of Riemannian curvature endomorphism, which quantifies the non-commutativity of second covariant derivative and is taken as a fundamental operator in describing the effects of curvature of a manifold. To end this section, the Section 4.5 describes some curvature-related concepts, such as tidal effects, geodesic deviation and sectional curvature.

4.1. Commutativity of Directional Derivatives in $\mathbb{R}^n$

Commutativity of directional derivatives is a consequence of Schwarz's theorem. Recall that, given a vector field $f : \mathbb{R}^n \to \mathbb{R}^m$, and a vector $v \in \mathbb{R}^n$, its directional derivative at a point p is defined as

$$D_p f(v) := \lim_{h\to 0}\frac{f(p + hv) - f(p)}{h}. \tag{152}$$

Given another direction $u \in \mathbb{R}^n$, one may compute the directional derivative of the vector field $p \mapsto D_p f(v)$ along u, namely

$$D_p(D_p(v))(u) := \lim_{k\to 0}\frac{D_{p+ku}f(v) - D_p f(v)}{k}. \tag{153}$$

In other terms, we have that

$$\begin{aligned}
&D_p(D_p(v))(u)\\
&= \lim_{k\to 0}\frac{1}{k}\left\{\lim_{h\to 0}\left(\frac{f(p + hv + ku) - f(p + ku)}{h} - \frac{f(p + hv) - f(p)}{h}\right)\right\}\\
&= \lim_{h\to 0}\frac{1}{h}\left\{\lim_{k\to 0}\left(\frac{f(p + ku + hv) - f(p + hv)}{k} - \frac{f(p + ku) - f(p)}{k}\right)\right\}\\
&= \lim_{h\to 0}\frac{D_{p+hv}f(u) - D_p f(u)}{h}\\
&= D_p(D_p(u))(v),
\end{aligned} \tag{154}$$

hence directional derivatives commute.

The above-discussed commutativity test for directional derivatives is based on the notion of iterated directional derivation. To extend such test to manifolds, it is necessary to discuss what it is meant by iterated covariant derivation and through which formulas such iterated derivatives may be computed.

4.2. Iterated Covariant Derivatives and Second Covariant Derivative

Covariant derivation inputs a vector field and outputs a further vector field. Hence, it is possible to compute the covariant derivative of a covariant derivative (and so fort) of a given vector field.

In particular, given a smooth curve γ with image in $\mathbb{M}$, which represents the trajectory generated by a dynamical system, it is conceivable to define the following vector fields along such trajectory:

$$
\begin{aligned}
\dot{\gamma} &= \boldsymbol{v} \text{ (velocity field)},\\
\nabla_t^\gamma \dot{\gamma} &= \boldsymbol{a} \text{ (covariant acceleration)},\\
\nabla_t^\gamma \nabla_t^\gamma \dot{\gamma} &= \boldsymbol{r} \text{ (covariant jolt/jerk)},\\
\nabla_t^\gamma \nabla_t^\gamma \nabla_t^\gamma \dot{\gamma} &= \boldsymbol{n} \text{ (covariant snap/jounce)},\\
(\nabla_t^\gamma)^4 \dot{\gamma} &= \boldsymbol{c} \text{ (covariant crackle/flounce)},\\
(\nabla_t^\gamma)^5 \dot{\gamma} &= \boldsymbol{p} \text{ (covariant pop/pounce)}.
\end{aligned}
\tag{155}
$$

(It is an interesting research exercise to look for a translation of these terms in one's local language. For the Italian readers, derivatives of a positional variable from the third order on are translated as *strappo, sbalzo, crepitìo, schiocco*).

In general, iterated covariant derivatives are better computed through the expression (114) which, applied to speed, reads

$$
\nabla_t^\gamma \dot{\gamma} = \ddot{\gamma} + \bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma}).
\tag{156}
$$

Let us consider an example of evaluation of covariant acceleration on the special orthogonal group.

Example 16. *On the special orthogonal group $\mathbb{SO}(n)$ endowed with its canonical metric, the covariant acceleration associated to a trajectory $t \to R(t)$ reads, according to Equation (156):*

$$
\boldsymbol{a}_t^R = \ddot{R} + \bar{\bar{\Gamma}}_R(\dot{R}, \dot{R}) = \ddot{R} + R\dot{R}^\top \dot{R}.
\tag{157}
$$

Recalling that $\mathbb{SO}(n)$ is a Lie group and writing the velocity $\dot{R}$ as RH, with $H \in \mathfrak{so}(n)$, one may write $\ddot{R} = \dot{R}H + R\dot{H} = R(H^2 + \dot{H})$, hence substituting:

$$
\boldsymbol{a}_t^R = \ddot{R} + R\dot{R}^\top \dot{R} = R(H^2 + \dot{H}) - RH^2 = R\dot{H}.
\tag{158}
$$

Whenever the group $\mathbb{SO}(3)$ is used to model the rotational dynamics of a rigid body, the above relation is interpreted as follows: The covariant acceleration equals the angular acceleration in the body-fixed reference frame brought back to an inertial reference frame.

Applying again the Equation (114) to the vector field $\ddot{\gamma} + \bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma})$ yields

$$
\nabla_t^\gamma \nabla_t^\gamma \dot{\gamma} = \dddot{\gamma} + \frac{\mathrm{d}}{\mathrm{d}t}\bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\ddot{\gamma} + \bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma}), \dot{\gamma}).
\tag{159}
$$

The same procedure may be repeated to get the expressions of the higher-order iterated covariant derivatives. It is worth observing that, for the first time, we encounter a derivative of the Christoffel form, $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\bar{\Gamma}}$, as well as nested Christoffel forms $\bar{\bar{\Gamma}}(\bar{\bar{\Gamma}}, \star)$. Such observation

turns out to hold in the more general setting of nested connections applied to vector fields and plays a fundamental role in the computation of *Riemannian curvature*.

It is worth underlying two special cases.

- A special case arises in the presence of a connection whose Christoffel form is normal, namely $\Pi_\gamma \bar{\bar{\Gamma}} \equiv 0$, in which case we may access the simplified relation (36) to compute covariant derivative, namely $\nabla_t^\gamma \dot{\gamma} = \Pi_\gamma\{\ddot{\gamma}\}$. In such a case, the covariant jolt may be computed as

$$\mathfrak{r}_t^\gamma = \Pi_\gamma\{\dddot{\gamma} + \Pi_\gamma^\bullet(\ddot{\gamma}, \dot{\gamma})\}. \tag{160}$$

- A even more particular case is given by a geodesic curve, for which $\ddot{\gamma} = -\bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma})$. In this case, the acceleration $\mathfrak{a}_t^\gamma = \Pi_\gamma\{\ddot{\gamma} + \bar{\bar{\Gamma}}_\gamma(\dot{\gamma}, \dot{\gamma})\} = 0$, hence all covariant derivatives (from acceleration to pounce) are identically zero (no matter what is the structure of the Cristoffel form of the second kind).

Let us examine a few examples of computation of covariant jolt $\mathfrak{r}_t^\gamma$ associated to a manifold-valued system trajectory.

Example 17. *The first case treated in the present example concerns the hypersphere $\mathbb{S}^{n-1}$ endowed with its canonical metric. In this case, the Christoffel form of the second kind reads $\bar{\bar{\Gamma}}_x(u, v) = (u^\top v)x$ hence the covariant acceleration reads*

$$\mathfrak{a}_t^x = \ddot{x} + x\dot{x}^\top \dot{x}. \tag{161}$$

In addition, according to Equation (159), the covariant jolt of a trajectory $t \mapsto x(t)$ in $\mathbb{S}^{n-1}$ reads

$$\begin{aligned}\mathfrak{r}_t^x &= \dddot{x} + \frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}^\top \dot{x}x) + \dot{x}(\ddot{x} + \dot{x}^\top \dot{x}x)x \\ &= \dddot{x} + 3\,x\ddot{x}^\top \dot{x} + \dot{x}\dot{x}^\top \dot{x}. \end{aligned} \tag{162}$$

Let us recall that from the condition $x^\top x = 1$, deriving thrice, one gets $3\,\ddot{x}^\top \dot{x} + x^\top \dddot{x} = 0$, hence the last expression is equivalent to

$$\mathfrak{r}_t^x = (I_n - xx^\top)\dddot{x} + \dot{x}\dot{x}^\top \dot{x}. \tag{163}$$

We might have obtained the same expression more easily recalling that, in this case, the Christoffel form is purely normal, hence the jolt is given by the Equation (160). Let us recall that $\Pi_x(a) = a - xx^\top a$, hence $\Pi_x^\bullet(a, \dot{x}) = -\dot{x}x^\top a - x\dot{x}^\top a$. Therefore, the jolt is computed as

$$\mathfrak{r}_t^x = (I - xx^\top)\dddot{x} - \dot{x}x^\top \ddot{x}. \tag{164}$$

Notice however that, deriving both sides of the condition $x^\top x = 1$ twice yields $x^\top \ddot{x} + \dot{x}^\top \dot{x} = 0$, hence the expression (164) is equivalent to the expression (163).

A second case-of-study arises from the special orthogonal group $\mathbb{SO}(n)$ endowed with its canonical metric. In this case, the jolt of a trajectory $t \mapsto R(t)$ in $\mathbb{SO}(n)$ computed through the general Equation (159), taking into account that $\bar{\bar{\Gamma}}_R(U, V) = \frac{1}{2}R(U^\top V + V^\top U)$, reads

$$\mathfrak{r}_t^R = \dddot{R} + \frac{\mathrm{d}}{\mathrm{d}t}(R\dot{R}^\top \dot{R}) + \frac{1}{2}R((\ddot{R} + R\dot{R}^\top \dot{R})^\top \dot{R} + \dot{R}^\top(\ddot{R} + R\dot{R}^\top \dot{R})). \tag{165}$$

Even in this case, the Christoffel form is purely normal, hence we may apply the relationship (160), upon recalling that $\Pi_R(A) = \frac{1}{2}A - \frac{1}{2}RA^\top R$ and hence $\Pi_R^\bullet(A, \dot{R}) = -\frac{1}{2}\dot{R}A^\top R - \frac{1}{2}RA^\top \dot{R}$, obtaining

$$\mathfrak{r}_t^R = \Pi_R\{\dddot{R} + \dot{R}\dot{R}^\top \dot{R}\}. \tag{166}$$

In this case, since the manifold $\mathbb{SO}(n)$ is a Lie group, the velocity over the trajectory may be factored as $\dot{R} = RH$ with $H \in \mathfrak{so}(n)$. Hence

$$
\begin{aligned}
\ddot{R} &= \dot{R}H + R\dot{H} = R(H^2 + \dot{H}), \\
\dddot{R} &= \dot{R}(H^2 + \dot{H}) + R(\dot{H}H + H\dot{H} + \ddot{H}) = R(\ddot{H} + 2H\dot{H} + \dot{H}H + H^3).
\end{aligned}
\tag{167}
$$

Therefore one may verify that the above expressions are equivalent to one another and to

$$
\mathbf{r}_t^R = R\ddot{H} + \tfrac{1}{2}R(H\dot{H} - \dot{H}H).
\tag{168}
$$

Notice that the result of the calculation is a tangent vector in $T_R\mathbb{SO}(n)$. The so-called rotational jolt $\mathbf{r}_t^R$ *(and, in particular, a function of its norm $\|\mathbf{r}_t^R\|_{R(t)}$) was used in several contexts to measure the fluidity of rotational movements. An example may be found in [20,21], which dealt with hips sway fluency assessment in the context of movement analysis.*

The last example considered concerns the space of symmetric, positive-definite matrices $\mathbb{S}^+(n)$ endowed with its canonical metric. In this case, $\bar{\bar{\Gamma}}_P(U,V) = -\tfrac{1}{2}UP^{-1}V - \tfrac{1}{2}VP^{-1}U$, hence

$$
\mathfrak{a}_t^P = \ddot{P} - \dot{P}P^{-1}\dot{P},
\tag{169}
$$

and, consequently,

$$
\begin{aligned}
\mathbf{r}_t^P &= \dddot{P} + \tfrac{d}{dt}(-\dot{P}P^{-1}\dot{P}) - \tfrac{1}{2}(\ddot{P} - \dot{P}P^{-1}\dot{P})P^{-1}\dot{P} - \tfrac{1}{2}\dot{P}P^{-1}(\ddot{P} - \dot{P}P^{-1}\dot{P}) \\
&= \dddot{P} - \tfrac{3}{2}(\ddot{P}P^{-1}\dot{P} + \dot{P}P^{-1}\ddot{P}) + 2\dot{P}P^{-1}\dot{P}P^{-1}\dot{P},
\end{aligned}
\tag{170}
$$

a (perhaps unexpectedly) convoluted expression. Notice that, in this case, the Christoffel form is not purely normal, hence the special relationship (160) does not hold.

We mention that, if a curve γ has unitary speed, namely $\|\dot{\gamma}\|_\gamma = 1$, then the norm of its covariant acceleration is taken as a measure of its curvature. Namely, the scalar field

$$
\kappa_g^\gamma := \|\nabla_t^\gamma \dot{\gamma}\|_\gamma,
\tag{171}
$$

is defined *geodesic curvature* and quantifies how much a curve differs from a geodesic. Hence a geodesic curve has zero geodesic curvature.

In a similar manner as we have done for the covariant derivatives along a curve, we may compute iterated derivatives of a given vector fields along different directions. Notationwise, given a vector field $w \in \Gamma(\mathbb{M})$ whose covariant derivative is sought and two directions $u, v \in T_x\mathbb{M}$, a nested derivative will be denoted as $\nabla_u \nabla_v w$. Let us make such abstract notation become substantial. The vector field w needs to be defined in a neighborhood of the point x. Likewise, it is necessary to set up a vector field $\bar{v} \in \Gamma(\mathbb{M})$, well-defined in a neighborhood of x, such that $\bar{v}_x = v$. In this way, the connection $\nabla_{\bar{v}}w \in \Gamma(\mathbb{M})$ yields a vector field, well-defined in a neighborhood of x, which may be further covariantly-derivated along the direction u.

In a even more natural manner, given three vector fields $u, v, w \in \Gamma(\mathbb{M})$, one may define an iterated connection $\nabla_u \nabla_v w$, whose result may be expressed as

$$
(\nabla_u \nabla_v w)_x = (\partial_x(\nabla_v w)_x)(u_x) + \bar{\bar{\Gamma}}_x((\nabla_v w)_x, u_x).
\tag{172}
$$

Notice that the term $\partial_x(\nabla_v w)(u)$ involves the directional derivative $\partial_x \bar{\bar{\Gamma}}_x$ of the Christoffel form, as well as the directional derivative of the vector fields w and v, in fact

$$
\begin{aligned}
(\partial_x(\nabla_v w)_x)(u_x) &= \partial_x(\partial_x w(v_x) + \bar{\bar{\Gamma}}_x(w_x, v_x))(u_x) \\
&= \partial_x^2 w(v_x; u_x) + (\partial_x w)((\partial_x v)(u_x)) + (\partial_x \bar{\bar{\Gamma}}_x)(w_x, v_x; u_x) + \\
&\quad \bar{\bar{\Gamma}}_x(\partial_x w(u_x), v_x) + \bar{\bar{\Gamma}}_x(w_x, \partial_x v(u_x)),
\end{aligned}
\tag{173}
$$

where $\partial_x^2 w(v_x; u_x)$ denotes the second directional derivative of the vector field w and $\partial_x(\bar{\bar{\Gamma}}_x)$ denotes the derivative of the Christoffel form with respect to the argument x only. Therefore, the nested connection may effectively be computed through the (a little convoluted) expression

$$
\begin{aligned}
(\nabla_u \nabla_v w)_x \\
= \partial_x^2 w(v_x; u_x) + (\partial_x w)((\partial_x v)(u_x)) + (\partial_x \bar{\bar{\Gamma}}_x)(w_x, v_x; u_x) + \\
\bar{\bar{\Gamma}}_x(\partial_x w(u_x), v_x) + \bar{\bar{\Gamma}}_x(w_x, \partial_x v(u_x)) + \bar{\bar{\Gamma}}_x(\partial_x w(v_x), u_x) + \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, v_x), u_x),
\end{aligned}
\tag{174}
$$

where we have made use repeatedly of the linearity of the Christoffel form of the second kind with respect to its vectorial arguments to extend it to the whole ambient space.

An unavoidable characteristic of the iterated derivative (174) is that it includes not only the rate of change of the vector field w and of the Christoffel form (which describes the underlying manifold) but also the rate of change of the vector field v, which may be an unwanted effect. Therefore, on the basis of the nested covariant derivative (174), one may define a *second covariant derivative* of the vector field w in the directions u, v as follows:

$$
\nabla_{u,v}^2 w := \nabla_u \nabla_v w - \nabla_{\nabla_u v} w.
\tag{175}
$$

To verify the effects of such definition, it pays to compute first

$$
\begin{aligned}
(\nabla_{\nabla_u v} w)_x &= (\partial_x w)((\nabla_u v)_x) + \bar{\bar{\Gamma}}_x(w_x, (\nabla_u v)_x) \\
&= (\partial_x w)(\partial_x v(u_x) + \bar{\bar{\Gamma}}_x(v_x, u_x)) + \bar{\bar{\Gamma}}_x(w_x, \partial_x v(u_x)) + \bar{\bar{\Gamma}}_x(w_x, \bar{\bar{\Gamma}}_x(v_x, u_x)),
\end{aligned}
\tag{176}
$$

so that the second covariant derivative may be expressed as

$$
\begin{aligned}
(\nabla_{u,v}^2 w)_x \\
= \partial_x^2 w(v_x; u_x) + (\partial_x w)((\partial_x v)(u_x)) + (\partial_x \bar{\bar{\Gamma}}_x)(w_x, v_x; u_x) \\
+ \bar{\bar{\Gamma}}_x(\partial_x w(u_x), v_x) + \bar{\bar{\Gamma}}_x(w_x, \partial_x v(u_x)) + \bar{\bar{\Gamma}}_x(\partial_x w(v_x), u_x) \\
+ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, v_x), u_x) - (\partial_x w)(\partial_x v(u_x)) - (\partial_x w)(\bar{\bar{\Gamma}}_x(v_x, u_x)) \\
- \bar{\bar{\Gamma}}_x(w_x, \partial_x v(u_x)) - \bar{\bar{\Gamma}}_x(w_x, \bar{\bar{\Gamma}}_x(v_x, u_x)) \\
= \partial_x^2 w(v_x; u_x) + (\partial_x \bar{\bar{\Gamma}}_x)(w_x, v_x; u_x) + \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, v_x), u_x) \\
+ \bar{\bar{\Gamma}}_x(\partial_x w(u_x), v_x) + \bar{\bar{\Gamma}}_x(\partial_x w(v_x), u_x) - (\partial_x w)(\bar{\bar{\Gamma}}_x(v_x, u_x)) \\
- \bar{\bar{\Gamma}}_x(w_x, \bar{\bar{\Gamma}}_x(v_x, u_x)).
\end{aligned}
\tag{177}
$$

The second derivative looks 'purified' from the rate of change of the vector field v.

Observation 1. *For the mathematically-minded, we underline that there exists a technical reason why the iterated covariant derivative $\nabla_u \nabla_v w$ is, in some sense, not well-defined while the double covariant derivative $\nabla_{u,v} w$ is well defined. Let us recall that the iterated covariant derivative $(\nabla_u \nabla_v w)$ is meant to capture the rate of change, along the direction u, of the rate of change of the vector field w along the direction v, at a given point x: in principle, hence $v \in T_x\mathbb{M}$ is a single tangent vector to which there need not to be associated any vector field. Nevertheless, in order to evaluate the covariant derivative along the direction u, it is necessary to 'extend' the direction v locally to a smooth vector field, say $\bar{v} \in \Gamma(\mathbb{M})$: such vector field needs only to obey the constraint $\bar{v}_x \equiv v$, being elsewhere unconstrained. Now, since the nested covariant derivative (174) depends explicitly on $\partial_x \bar{v}$, its value depends on the way we have extended the vector v by the field $\bar{v}$, which makes its expression 'user-dependent'. As opposed to this, the second covariant derivative (177) does not exhibit such kind of dependence, namely the way one chooses the extension $\bar{v}$ is irrelevant to the value of the second covariant derivative, which results henceforth well-defined ('user-independent').*

Now, we may wonder whether the second covariant derivative, which depends on the second directional derivative, is independent of the order of derivation, namely if the result

changes if one swaps the vectors u and v. From the last expression, we may immediately find that

$$
\begin{aligned}
(\nabla^2_{v,u}w)_x &= \partial^2_x w(u_x; v_x) + (\partial_x \bar{\bar{\Gamma}}_x)(w_x, u_x; v_x) + \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, u_x), v_x) \\
&\quad + \bar{\bar{\Gamma}}_x(\partial_x w(v_x), u_x) + \bar{\bar{\Gamma}}_x(\partial_x w(u_x), v_x) - (\partial_x w)(\bar{\bar{\Gamma}}_x(u_x, v_x)) \\
&\quad - \bar{\bar{\Gamma}}_x(w_x, \bar{\bar{\Gamma}}_x(u_x, v_x)).
\end{aligned}
\tag{178}
$$

By exploiting the known symmetry of the Christoffel form along with evident cancellations, one immediately finds that

$$
\begin{aligned}
(\nabla^2_{u,v}w)_x - (\nabla^2_{v,u}w)_x &= (\partial_x \bar{\bar{\Gamma}}_x)(w_x, v_x; u_x) - (\partial_x \bar{\bar{\Gamma}}_x)(w_x, u_x; v_x) + \\
&\quad \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, v_x), u_x) - \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, u_x), v_x).
\end{aligned}
\tag{179}
$$

The net result is that *the second covariant derivative does not commute* and *the degree of non-commutativity depends solely on the behavior of the Christoffel form of the second kind* while it does not depend at all by the rate of change of the vector field being derivated (w). A remarkable consequence of the latter observation is that the quantity $(\nabla^2_{u,v}w)_x - (\nabla^2_{v,u}w)_x$ *is an intrinsic property of manifolds*.

Let us underline that, for a torsion-free connection as the Levi-Civita one, the non-commutativity measure $\nabla^2_{u,v} - \nabla^2_{v,u}$ may be written in terms of commutators. In fact, from the definition of second covariant derivative (175) we get that

$$
\begin{aligned}
\nabla^2_{u,v}w - \nabla^2_{v,u}w &= \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{\nabla_u v}w + \nabla_{\nabla_v u}w \\
&= (\nabla_u \nabla_v - \nabla_v \nabla_u - \nabla_{\nabla_u v - \nabla_v u})w.
\end{aligned}
\tag{180}
$$

Now, recalling that $\nabla_u v - \nabla_v u = [u, v]$ and defining the following covariant differential commutator $[\nabla_u, \nabla_v] := \nabla_u \nabla_v - \nabla_v \nabla_u$, the above relationship may be rewritten compactly and meaningfully as

$$
\nabla^2_{u,v} - \nabla^2_{v,u} = [\nabla_u, \nabla_v] - \nabla_{[u,v]},
\tag{181}
$$

which suggests that the lack of commutativity of the iterated covariant connection is due to the lack of commutativity of covariant connection depurated from the lack of commutativity of the involved vector fields.

To conclude this section, let us examine, as an example, how to generalize the relationship (159) to a generic tangent vector field.

Example 18. *Given a smooth vector field $w \in \Gamma(\mathbb{M})$, its covariant derivative along a smooth curve γ may be expressed through the relation (114) as $\nabla^\gamma_t w = \dot{w}_\gamma + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma})$. Applying again the Equation (114) to the vector field $\dot{w}_\gamma + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma})$ yields*

$$
\begin{aligned}
\nabla^\gamma_t \nabla^\gamma_t w &= \ddot{w}_\gamma + \frac{\mathrm{d}}{\mathrm{d}t} \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\dot{w}_\gamma + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}), \dot{\gamma}) \\
&= \ddot{w}_\gamma + \bar{\bar{\Gamma}}^\bullet_\gamma(w_\gamma, \dot{\gamma}; \dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\dot{w}_\gamma, \dot{\gamma}) \\
&\quad + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \ddot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\dot{w}_\gamma, \dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}), \dot{\gamma}) \\
&= \ddot{w}_\gamma + \bar{\bar{\Gamma}}^\bullet_\gamma(w_\gamma, \dot{\gamma}; \dot{\gamma}) + 2\bar{\bar{\Gamma}}_\gamma(\dot{w}_\gamma, \dot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \ddot{\gamma}) + \bar{\bar{\Gamma}}_\gamma(\bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}), \dot{\gamma}),
\end{aligned}
\tag{182}
$$

where the following abbreviation has been made use of

$$
\bar{\bar{\Gamma}}^\bullet_\gamma(a, b; \dot{\gamma}) := \frac{\mathrm{d}}{\mathrm{d}t} \bar{\bar{\Gamma}}_\gamma(a, b),
\tag{183}
$$

with $a, b \in \mathbb{A}$ being arbitrary constants. As we shall see in the next section of this paper, such linear operator is not just a convenient placeholder but plays indeed a fundamental role in quantifying the Riemannian notion of curvature.

4.3. Manifestation of Manifold Curvature in Parallel Transport along a Loop

Let ℓ denote a simple loop, namely a simple curve on a manifold that starts and ends at the same point x and let $w_x \in T_x\mathbb{M}$ denote a test tangent vector. Let us parallely-transport the tangent vector w_x along the closed loop ℓ and let us denote the final result of transport as $w_\ell \in T_x\mathbb{M}$.

If a manifold is flat, the result of parallel transport along a loop is just the initial vector, namely $w_\ell \equiv w_x$. On a curved manifold, *this is not necessarily the case*, as a manifestation of curvedness. Let us examine such phenomenon, referred to as *anholonomy*, in detail, in the hypothesis that the closed loop ℓ be very small.

To effect parallel transport along the loop ℓ parameterized by $\gamma : [0, 1] \to \mathbb{M}$, with $\gamma(0) = x$, we should set-up the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}w_\gamma + \bar{\bar{\Gamma}}_\gamma(w_\gamma, \dot{\gamma}) = 0 \tag{184}$$

in the transport field $w \in \Gamma(\mathbb{M})$, with initial condition $w_{\gamma(0)} = w_x$ given. Let us observe that, according to the property of the Christoffel form underlined in Section 3.8, the solution of the above differential equation in $\mathbb{A}$ is indeed a tangent vector field along γ, hence we may write

$$w_{\gamma(t)} = w_x - \int_0^t \bar{\bar{\Gamma}}_{\gamma(\tau)}(w_{\gamma(\tau)}, \dot{\gamma}(\tau))\mathrm{d}\tau. \tag{185}$$

Setting $t = 1$ leads to

$$w_\ell = w_x - \oint_\ell \bar{\bar{\Gamma}}_{\gamma(\tau)}(w_{\gamma(\tau)}, \dot{\gamma}(\tau))\mathrm{d}\tau, \tag{186}$$

since the loop is closed.

Under the hypothesis that the loop is very small, it makes sense to approximate the loop integral above by means of first-order approximations of the Christoffel form as well as the transport field around the foot-point x, namely

$$\bar{\bar{\Gamma}}_{\gamma(\tau)}(a, b) = \bar{\bar{\Gamma}}_x(a, b) + \bar{\bar{\Gamma}}_x^\bullet(a, b; \gamma(\tau) - x) + \mathscr{O}(\tau^2), \tag{187}$$

$$w_{\gamma(\tau)} = w_x + \tau\dot{w}_x + \mathscr{O}(\tau^2) = w_x - \bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x) + \mathscr{O}(\tau^2), \tag{188}$$

with $\tau \in [0, 1]$, where we have made use of the definition (183) and the last equation derives from the relationship (184). Also, we have made use of the approximation $\dot{\gamma}(\tau) = \frac{\gamma(\tau)-x}{\tau} + \mathscr{O}(\tau)$. On the basis of the above approximations, we see that

$$\begin{aligned}
\bar{\bar{\Gamma}}_{\gamma(\tau)}(w_{\gamma(\tau)}, \dot{\gamma}(\tau)) &= \bar{\bar{\Gamma}}_x(w_x - \bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x) + \mathscr{O}(\tau^2), \dot{\gamma}(\tau)) + \\
&\quad \bar{\bar{\Gamma}}_x^\bullet(w_x - \bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x) + \mathscr{O}(\tau^2), \dot{\gamma}(\tau); \gamma(\tau) - x) + \mathscr{O}(\tau^2) \\
&= \bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}(\tau)) - \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau)) + \\
&\quad \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}(\tau); \gamma(\tau) - x) - \bar{\bar{\Gamma}}_x^\bullet(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau); \gamma(\tau) - x) \\
&\quad + \mathscr{O}(\tau^2),
\end{aligned} \tag{189}$$

therefore, making use of the approximation $\bar{\bar{\Gamma}}_x^\bullet(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau); \gamma(\tau) - x) = \mathscr{O}(\tau^2)$, we get

$$\begin{aligned}
\bar{\bar{\Gamma}}_{\gamma(\tau)}(w_{\gamma(\tau)}, \dot{\gamma}(\tau)) &= \bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}(\tau)) - \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau)) + \\
&\quad \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}(\tau); \gamma(\tau) - x) + \mathscr{O}(\tau^2).
\end{aligned} \tag{190}$$

Plugging the above relation into the loop integral (186), we get

$$\begin{aligned}
w_\ell \approx w_x - \oint_\ell \bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}(\tau))\mathrm{d}\tau + \oint_\ell \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau))\mathrm{d}\tau \\
- \oint_\ell \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}(\tau); \gamma(\tau) - x)\mathrm{d}\tau.
\end{aligned} \tag{191}$$

Recalling that the Christoffel form is (bi)linear in its vectorial arguments, it is immediate to see that the first addendum is null, namely

$$\oint_\ell \bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}(\tau))\mathrm{d}\tau = \bar{\bar{\Gamma}}_x\left(w_x, \oint_\ell \dot{\gamma}(\tau)\mathrm{d}\tau\right) = \bar{\bar{\Gamma}}_x(w_x, 0) = 0, \tag{192}$$

therefore ultimately we get that the discrepancy between the transported vector after and before transport is quantified by

$$w_\ell - w_x \approx \oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau) - x), \dot{\gamma}(\tau)) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}(\tau); \gamma(\tau) - x) \right\}\mathrm{d}\tau \tag{193}$$

up to order $\mathcal{O}(\tau^3)$. The above relationship may be written in a more 'symmetric' way by the following arguments. Let us first notice that

$$\oint_\ell \frac{\mathrm{d}}{\mathrm{d}\tau}\left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma(\tau)), \gamma(\tau)) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \gamma(\tau); \gamma(\tau)) \right\}\mathrm{d}\tau = 0, \tag{194}$$

therefore it holds that

$$\oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma), \dot{\gamma}) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}; \gamma) \right\}\mathrm{d}\tau + \oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}), \gamma) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \gamma; \dot{\gamma}) \right\}\mathrm{d}\tau = 0. \tag{195}$$

In addition, notice that, by linearity and closedness of the integration path, it holds that

$$\oint_\ell \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, x), \dot{\gamma})\mathrm{d}\tau = \oint_\ell \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}), x)\mathrm{d}\tau = 0, \tag{196}$$

$$\oint_\ell \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}; x)\mathrm{d}\tau = \oint_\ell \bar{\bar{\Gamma}}_x^\bullet(w_x, x; \dot{\gamma})\mathrm{d}\tau = 0, \tag{197}$$

therefore, the relationship (195) may be rewritten as

$$\begin{aligned}
\oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma - x), \dot{\gamma}) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}; \gamma - x) \right\}\mathrm{d}\tau = \\
- \oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}), \gamma - x) - \bar{\bar{\Gamma}}_x^\bullet(w_x, \gamma - x; \dot{\gamma}) \right\}\mathrm{d}\tau.
\end{aligned} \tag{198}$$

Through this equation, the discrepancy (193) may be rewritten as

$$\begin{aligned}
w_\ell - w_x \approx \frac{1}{2}\oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \gamma - x), \dot{\gamma}) + \bar{\bar{\Gamma}}_x^\bullet(w_x, \gamma - x; \dot{\gamma}) \right\}\mathrm{d}\tau \\
- \frac{1}{2}\oint_\ell \left\{ \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w_x, \dot{\gamma}), \gamma - x) + \bar{\bar{\Gamma}}_x^\bullet(w_x, \dot{\gamma}; \gamma - x) \right\}\mathrm{d}\tau.
\end{aligned} \tag{199}$$

One may hence infer that the integrand is hence responsible for the discrepancy between w_ℓ and w_x. In addition, we notice informally that the integrand appears to have exactly the same structure of the right-hand side of the relation (179). (In addition, from the above relation it is easily guessed that the amount of discrepancy increases with the size of the loop.)

4.4. Riemannian Curvature Endomorphism

As we have seen, given a closed loop ℓ with foot-point $x \in \mathbb{M}$ and a tangent vector $w_x \in T_x\mathbb{M}$, the parallel transport of the latter along the closed loop results in a tangent vector $w_\ell \in T_x\mathbb{M}$ that differs from w_x. Since, however, both vectors w_x and w_ℓ belong to the same tangent space (irrespective of how ample is the loop), one may postulate the existence of a function, call it $\mathbf{R}_x : T\mathbb{M} \to T\mathbb{M}$, that transforms w_x to w_ℓ. Such function is called *curvature endomorphism* and we write $w_\ell = \mathbf{R}_x w_x$.

As it is customary, we are going to determine the structure of the curvature endomorphism for a special loop, namely a *parallelogram*. In order to define a parallelogram on a smooth manifolds' tangent bundle, let us define the following elements

- four points $x, a, b, c \in \mathbb{M}$ that denote the vertexes of the parallelogram in $\mathbb{M}$ around the point x,
- two tangent vectors $u, v \in T_x\mathbb{M}$ that define the 'sides' of the parallelogram,
- a smooth curve $\alpha : [0, 1] \to \mathbb{M}$ that joins the point x to the point a, such that $\alpha(0) = x$ and $\dot{\alpha}(0) = u$; the curve α is parametrized through the parameter s, namely a point on such curve is denoted by $\alpha(s)$,
- a smooth curve $\beta : [0, 1] \to \mathbb{M}$ that joins the point a to the point c, such that $\beta(0) = \alpha(s)$ and $\dot{\beta}(0) = \mathbb{P}_\alpha^{0 \to s}(v)$; the curve β is parametrized through the parameter t, namely a point on such curve is denoted by $\beta(t)$,
- a smooth curve $\lambda : [0, 1] \to \mathbb{M}$ that joins the point x to the point b, such that $\lambda(0) = x$ and $\dot{\lambda}(0) = v$; the curve λ is parametrized through the parameter t, namely a point on such curve is denoted by $\lambda(t)$,
- a smooth curve $\gamma : [0, 1] \to \mathbb{M}$ that joins the point b to the point b, such that $\gamma(0) = \lambda(t)$ and $\dot{\gamma}(0) = \mathbb{P}_\lambda^{0 \to t}(u)$; the curve γ is parametrized through the parameter s, namely a point on such curve is denoted by $\gamma(s)$.

Notice that, by consistency, the point $\beta(t)$ on the parallelogram coincides to the point $\gamma(s)$ which, in turn, coincides to the point c, which lays at the 'opposite' side to the point x in such parallelogram. Also, notice that the extension of the parallelogram depends on the values of the parameters s and t, hence such parallelogram may be made as small as one wishes by reducing the values of such parameters.

It will be necessary to make use of local extensions of the vectors u, v and of the test tangent vector $w \in T_x\mathbb{M}$ being parallely-transported. In order not to burden the notation with extra symbols, we shall denote such (arbitrary) extensions by $U, V, W \in \Gamma(\mathbb{M})$, respectively. These vector fields may be chosen arbitrarily as long as they are sufficiently regular to afford derivation and $U_x = u, V_x = v, W_x = w$.

As a further introductory element, let us recall a couple of very useful first-order approximations. Let $\rho : [-\epsilon, \epsilon] \to \mathbb{M}$ denote a smooth curve in $\mathbb{M}$ and $F \in \Gamma(\mathbb{M})$ denote a smooth vector field. From the definition of covariant derivative of a vector field along a given direction we get the following first-order approximation of the vector field F at $\rho(0)$:

$$(\textbf{A1}): \quad \mathbb{P}_\rho^{\tau \to 0} F_{\rho(\tau)} = F_{\rho(0)} + \tau \, \nabla_{\dot{\rho}(0)} F + \mathcal{O}(\tau^2). \tag{200}$$

This relationship may be rewritten equivalently as $F_{\rho(\tau)} = \mathbb{P}_\rho^{0 \to \tau} F_{\rho(0)} + \tau \, \mathbb{P}_\rho^{0 \to \tau}(\nabla_{\dot{\rho}(0)} F) + \mathcal{O}(\tau^2)$, which appears as a sort of Taylor expansion of a vector field along a curve departing from a given point. Such expression leads to the further approximation

$$(\textbf{A2}): \quad \mathbb{P}_\rho^{0 \to \tau} F_{\rho(0)} = F_{\rho(\tau)} - \tau \, \mathbb{P}_\rho^{0 \to \tau}(\nabla_{\dot{\rho}(0)} F) + \mathcal{O}(\tau^2). \tag{201}$$

The approximations $(\textbf{A1})$ and $(\textbf{A2})$ will prove instrumental in writing down the following necessary relations seamlessly.

The main idea to define the curvature tensor is that transporting a tangent vector along a parallelogram and comparing the starting vector with the resulting vector is equivalent to transporting the same vectors along two paths, namely along $x \to a \to c$ and along $x \to b \to c$ and then comparing the two results at c. Formally, we hence define the Riemannian curvature endomorphism as

$$\mathbf{R}_x^{u,v}(w) := \lim_{s,t \to 0} \frac{\mathbb{P}_\alpha^{s \to 0} \, \mathbb{P}_\beta^{t \to 0} \, W_{\beta(t)} - \mathbb{P}_\lambda^{t \to 0} \, \mathbb{P}_\gamma^{s \to 0} \, W_{\gamma(t)}}{s \, t}. \tag{202}$$

For the first branch, we have

$$
\begin{aligned}
\mathbb{P}_\alpha^{s\to0}\,\mathbb{P}_\beta^{t\to0}\,W_{\beta(t)} &= \mathbb{P}_\alpha^{s\to0}\big(W_{\beta(0)} + t\,\nabla_{\dot{\beta}(0)}W + \mathscr{O}(t^2)\big)\\
&= \mathbb{P}_\alpha^{s\to0}W_{\alpha(s)} + t\,\mathbb{P}_\alpha^{s\to0}\big(\nabla_{\mathbb{P}_\alpha^{0\to s}(v)}W\big) + \mathscr{O}(t^2)\\
&= W_{\alpha(0)} + s\,\nabla_{\dot{\alpha}(0)}W + \mathscr{O}(s^2) + t\,\mathbb{P}_\alpha^{s\to0}\big(\nabla_{V_{\alpha(s)} - s\,\mathbb{P}_\alpha^{0\to s}(\nabla_{\dot{\alpha}(0)}V) + \mathscr{O}(s^2)}W\big)\\
&\quad + \mathscr{O}(t^2)\\
&= w + s\,\nabla_u W + \mathscr{O}(s^2) + t\,\mathbb{P}_\alpha^{s\to0}\big(\nabla_{V_{\alpha(s)}}W - s\,\nabla_{\mathbb{P}_\alpha^{0\to s}(\nabla_u V)}W\big) + \mathscr{O}(t^2),
\end{aligned}
$$

where the first equality follows from the relation **A1** and the third equality follows from the relations **A1** and **A2**. As a consequence, it holds that

$$
\begin{aligned}
\mathbb{P}_\alpha^{s\to0}\,\mathbb{P}_\beta^{t\to0}\,W_{\beta(t)} &= w + s\,\nabla_u W + t\,\mathbb{P}_\alpha^{s\to0}\big(\nabla_{V_{\alpha(s)}}W\big) - s\,t\,\mathbb{P}_\alpha^{s\to0}\big(\nabla_{\mathbb{P}_\alpha^{0\to s}(\nabla_u V)}W\big)\\
&\quad + \mathscr{O}(s^2) + \mathscr{O}(t^2).
\end{aligned}
\tag{203}
$$

Analogously, for the second branch we get

$$
\begin{aligned}
\mathbb{P}_\lambda^{t\to0}\,\mathbb{P}_\gamma^{s\to0}\,W_{\gamma(t)} &= w + t\,\nabla_v W + s\,\mathbb{P}_\lambda^{t\to0}\big(\nabla_{U_{\lambda(t)}}W\big) - s\,t\,\mathbb{P}_\lambda^{t\to0}\big(\nabla_{\mathbb{P}_\lambda^{0\to t}(\nabla_v U)}W\big)\\
&\quad + \mathscr{O}(s^2) + \mathscr{O}(t^2).
\end{aligned}
\tag{204}
$$

Hence, after canceling out equal terms and grouping, we get

$$
\begin{aligned}
\mathbb{P}_\alpha^{s\to0}\,\mathbb{P}_\beta^{t\to0}\,W_{\beta(t)} &- \mathbb{P}_\lambda^{t\to0}\,\mathbb{P}_\gamma^{s\to0}\,W_{\gamma(t)} =\\
t\,\big\{\mathbb{P}_\alpha^{s\to0}&(\nabla_{V_{\alpha(s)}}W) - \nabla_v W\big\} - s\,\big\{\mathbb{P}_\lambda^{t\to0}(\nabla_{U_{\lambda(t)}}W) - \nabla_u W\big\}\\
- s\,t\,\big\{&\mathbb{P}_\alpha^{s\to0}(\nabla_{\mathbb{P}_\alpha^{0\to s}(\nabla_u V)}W) - \mathbb{P}_\lambda^{t\to0}(\nabla_{\mathbb{P}_\lambda^{0\to t}(\nabla_v U)}W)\big\} + \mathscr{O}(s^2) + \mathscr{O}(t^2).
\end{aligned}
\tag{205}
$$

Dividing both sides by the product $s\,t$ and taking the limit for s, t that vanish to zero, we get

$$
\begin{aligned}
\lim_{s,t\to0}\frac{\mathbb{P}_\alpha^{s\to0}\,\mathbb{P}_\beta^{t\to0}\,W_{\beta(t)} - \mathbb{P}_\lambda^{t\to0}\,\mathbb{P}_\gamma^{s\to0}\,W_{\gamma(t)}}{s\,t} =&\\
\lim_{s\to0}\frac{\mathbb{P}_\alpha^{s\to0}(\nabla_{V_{\alpha(s)}}W) - \nabla_v W}{s} - \lim_{t\to0}&\frac{\mathbb{P}_\lambda^{t\to0}(\nabla_{U_{\lambda(t)}}W) - \nabla_u W}{t}\\
- (\nabla_{\nabla_u V}W - \nabla_{\nabla_v U}W).&
\end{aligned}
\tag{206}
$$

In conclusion, upon computing the above limits, we get the expression of the curvature endomorphism applied to a tangent vector as follows

$$
\begin{aligned}
\mathbf{R}_x^{u,v}(w) &= (\nabla_u\nabla_V W)_x - (\nabla_v\nabla_U W)_x - (\nabla_{\nabla_u V - \nabla_v U}W)_x,\\
&= (\nabla_u\nabla_V W)_x - (\nabla_v\nabla_U W)_x - (\nabla_{[U,V]_x}W)_x.
\end{aligned}
\tag{207}
$$

At a first glance, it is not apparent that the combination of differential operators on the right-hand side of Equation (207) should equal a linear endomorphism as in the left-hand side. However, from Section 4.2 we know that not only this is indeed the case, but also that the right-hand side is completely independent from the way we choose the extensions U, V, W. In addition, from Section 4.3, we know that the expression on the right-hand side is exactly the combination of operators $\bar{\bar{\Gamma}}^\bullet$ and $\bar{\bar{\Gamma}}(\bar{\bar{\Gamma}}, \star)$ that quantifies the curvedness of a manifold in the parallel transport around a tiny loop. As a matter of fact, the Riemannian curvature endomorphism, applied to a test tangent vector $w \in T_x\mathbb{M}$, may be written explicitly (and solely) in terms of the Christoffel form of the second kind, as

$$
\mathbf{R}_x^{u,v}(w) = \bar{\bar{\Gamma}}_x^\bullet(w,v;u) - \bar{\bar{\Gamma}}_x^\bullet(w,u;v) + \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w,v),u) - \bar{\bar{\Gamma}}_x(\bar{\bar{\Gamma}}_x(w,u),v).
\tag{208}
$$

Let us apply the above formula to compute the curvature of a few manifolds of interest.

Example 19. *The first case examined is that of a unit hyper-sphere endowed with its canonical metric. Let us recall that $\bar{\bar{\Gamma}}_x(u,v) = x(u^\top v)$, hence*

$$\bar{\bar{\Gamma}}_x^\bullet(u,v;w) = w(u^\top v). \tag{209}$$

Applying the Equation ([208]) yields

$$\mathbf{R}_x^{u,v}(w) = u(w^\top v) - v(w^\top u) + x((x(w^\top v))^\top u) - x((x(w^\top u))^\top v). \tag{210}$$

Since the last two addenda are null because $x^\top u = x^\top v = 0$, we rest with

$$\mathbf{R}_x^{u,v}(w) = (w^\top v)u - (w^\top u)v. \tag{211}$$

Notice that the result does not depend explicitly on the point x, hence such Riemannian curvature endomorphism is uniform (in the sense that it takes the same expression in all tangent spaces). It is interesting to observe that the vector $\mathbf{R}_x^{u,v}(w)$ belongs to the subspace $\mathrm{span}\{u,v\} \subset T_x\mathbb{M}$.

Next we tackle the case of the special orthogonal group $\mathbb{SO}(n)$ endowed with its canonical metric. Let us recall again that the Christoffel form of the second kind reads $\bar{\bar{\Gamma}}_R(U,V) = \frac{1}{2}R(U^\top V + V^\top U)$, therefore

$$\bar{\bar{\Gamma}}_R^\bullet(U,V;W) = \frac{1}{2}W(U^\top V + V^\top U). \tag{212}$$

Applying the Equation ([208]) gives

$$\begin{aligned}
\mathbf{R}_R^{U,V}(W) =\ & \tfrac{1}{2}U(W^\top V + V^\top W) - \tfrac{1}{2}V(W^\top U + U^\top W) \\
& + \tfrac{1}{2}R\{\tfrac{1}{2}(W^\top V + V^\top W)^\top R^\top U + \tfrac{1}{2}U^\top R(W^\top V + V^\top W)\} \\
& - \tfrac{1}{2}R\{\tfrac{1}{2}(W^\top U + U^\top W)^\top R^\top V + \tfrac{1}{2}V^\top R(W^\top U + U^\top W)\} \\
=\ & \tfrac{1}{2}UW^\top V + \tfrac{1}{2}UV^\top W - \tfrac{1}{2}VW^\top U - \tfrac{1}{2}VU^\top W \\
& + \tfrac{1}{4}RW^\top VR^\top U + \tfrac{1}{4}RV^\top WR^\top U + \tfrac{1}{4}RU^\top RW^\top V + \tfrac{1}{4}RU^\top RV^\top W \\
& - \tfrac{1}{4}RW^\top UR^\top V - \tfrac{1}{4}RU^\top WR^\top V - \tfrac{1}{4}RV^\top RW^\top U - \tfrac{1}{4}RV^\top RU^\top W.
\end{aligned} \tag{213}$$

The addenda in the last expression do not look like they are going to add up and cancel nicely. Easy but lengthy calculations show that, indeed, they do. We are going to show only a few of the many transformations that are necessary:

$$\begin{aligned}
RW^\top(VR^\top)U &= -R(W^\top R)V^\top U = (RR^\top)WV^\top U = WV^\top U, \\
RV^\top(WR^\top)U &= -R(V^\top R)W^\top U = (RR^\top)VW^\top U = VW^\top U \\
(RU^\top)RW^\top V &= -U(R^\top R)W^\top V = -UW^\top V.
\end{aligned}$$

The net result is

$$\mathbf{R}_R^{U,V}(W) = \frac{1}{4}\left\{(UV^\top - VU^\top)W + W(V^\top U - U^\top V)\right\} \tag{214}$$

By some matrix manipulations similar to the above and by introducing the $T\mathbb{SO}(n)$-commutator $[U,V]_R := R[R^\top U, R^\top V]$ based on the matrix commutator $[A,B] := AB - BA$, the above relationship may be rewritten as

$$\mathbf{R}_R^{U,V}(W) = -\frac{1}{4}[[U,V]_R, W]_R. \tag{215}$$

Such expression is precisely the one reported in [22].

As a last example, let us consider the space of symmetric, positive-definite matrices $\mathbb{S}^{+}(n)$ endowed with its canonical metric, for which we had found that $\bar{\Gamma}_{P}(U,V) = -\frac{1}{2}UP^{-1}V - \frac{1}{2}VP^{-1}U$. The directional derivative of the Christoffel form reads

$$\bar{\bar{\Gamma}}_{P}^{\bullet}(U,V;W) = \tfrac{1}{2}UP^{-1}WP^{-1}V + \tfrac{1}{2}VP^{-1}WP^{-1}U. \tag{216}$$

By applying the Equation (208) we obtain an expression that consists of 16 terms, which however either cancel or add up nicely to one another, resulting in the final expression

$$\mathbf{R}_{P}^{U,V}(W) = \tfrac{1}{4}(VP^{-1}U - UP^{-1}V)P^{-1}W + \tfrac{1}{4}WP^{-1}(UP^{-1}V - VP^{-1}U), \tag{217}$$

that coincides to the one reported in [22].

The Riemannian curvature endomorphism enjoys, together with linearity in its vectorial argument, a number of symmetries. We are going to recall here only those that we are going to actually utilize in Section 4.5, namely

$$\begin{aligned} \langle \mathbf{R}^{u,v}(w), z \rangle &= -\langle \mathbf{R}^{v,u}(w), z \rangle, \\ \langle \mathbf{R}^{u,v}(w), z \rangle &= -\langle \mathbf{R}^{u,v}(z), w \rangle. \end{aligned} \tag{218}$$

In particular, one might notice that $\langle \mathbf{R}^{v,v}(\star), \star \rangle = 0$ and $\langle \mathbf{R}^{\star,\star}(v), v \rangle = 0$.

4.5. Tidal Effects, Geodesic Deviation, Jacobi Fields, Sectional Curvature

A massive body subjected to a spatially-varying gravitation field is subjected to tidal effects. In fact, since different parts of the body are subjected to a gravitation pull that varies from point to point, the whole body is subjected to internal mechanical stress.

Formally, such physical effect has an analogous in manifold calculus in the separation of nearby geodesics that may accelerate or decelerate according to the local curvature of a manifold. To quantify such effect, let us define a smooth family of geodesics $\gamma_{s}(t)$ where

- for a fixed value of the index $s \in [-\epsilon, \epsilon]$, the curve $\gamma_{s}(t)$ represents a geodesic parametrized by the parameter $t \in [0, 1]$; the geodesic curve $\gamma_{0}(t)$ represents the fiducial geodesic; each tangent vector $v_{s}(t) := \frac{\partial \gamma_{s}(t)}{\partial t}$ denotes the speed of the geodesic (we shall omit the values of the parameters for the sake of notation conciseness); all geodesics depart from the same point $x \in \mathbb{M}$, namely $\gamma_{s}(0) = x$ for any s;
- for a fixed value of the parameter $t \in [0, 1]$, the function $\gamma_{s}(t)$ represents a smooth curve (but not a geodesic); each tangent vector $\xi_{s}(t) := \frac{\partial \gamma_{s}(t)}{\partial s}$ represents the separation between two nearby geodesics (even in this case, we shall omit the values of the parameters for the sake of notation conciseness); more intuitively, the infinitesimal separation between nearby geodesics whose index s differ by an infinitesimal $\mathrm{d}s$ is $\xi\,\mathrm{d}s$. Since all geodesics depart from the same point, it holds that $\xi_{s}(0) = 0$ for every value of the index s.

Let us observe that the vector fields v and ξ commute, in fact

$$[v, \xi] = \frac{\partial^{2}}{\partial t \partial s}\gamma_{s}(t) - \frac{\partial^{2}}{\partial s \partial t}\gamma_{s}(t) = 0 \tag{219}$$

by the Schwarz-Clairaut's theorem.

The vector ξ may be referred to as *geodesic deviation*. It turns out that such deviations obeys a very neat equation involving the curvature endomorphism. The velocity with which the separation ξ between two nearby geodesic changes is quantified by the covariant derivation $\nabla_{v}\xi$, while the acceleration of the geodesic deviation is quantified by $\nabla_{v}\nabla_{v}\xi$. Let us recall that, for a torsion-free connection is holds that $\nabla_{v}\xi - \nabla_{\xi}v = [v, \xi]$. Since $[v, \xi] = 0$, it follows that $\nabla_{v}\xi = \nabla_{\xi}v$, hence

$$\nabla_{v}\nabla_{v}\xi = \nabla_{v}\nabla_{\xi}v. \tag{220}$$

From the Equation (207) it follows that

$$\nabla_v(\nabla_\xi v) - \nabla_\xi(\nabla_v v) - \nabla_{[v,\xi]}v = \mathbf{R}^{v,\xi}(v), \tag{221}$$

where we omitted the base-point x. Since $[v,\xi] = 0$ and $\nabla_v v = 0$ because v is the velocity vector of a geodesic, it ultimately follows that

$$\nabla_v\nabla_v\xi = \mathbf{R}^{v,\xi}(v), \tag{222}$$

that confirms that the acceleration with which two nearby geodesic originating from the same point deviate from one another depends on the local curvature of the underlying manifold.

The Equation (222) is called a *Jacobi equation* and its solution ξ is termed Jacobi field. In some circumstances, the Jacobi equation may assume a particularly simple form. To survey one of such cases, let us introduce the notion of *sectional curvature*.

The curvature endomorphism describes completely the curvature of a manifold in a given point, however the local curvature of a manifold is more easily described by a scalar index, the *sectional curvature* $K : (T\mathbb{M})^2 \to \mathbb{R}$, defined as

$$K_x^{u,v} := \frac{\langle \mathbf{R}_x^{u,v}(v), u \rangle_x}{\langle u, u \rangle_x \langle v, v \rangle_x - \langle u, v \rangle_x^2}. \tag{223}$$

Although the sectional curvature returns a scalar value, it by no means carries any diminished information about the curvature of a manifold, in fact it is possible to recover the curvature endomorphism by the sectional curvature. A special case that serves the purpose of illustrating such property is that of spaces of constant sectional curvature. A manifold $\mathbb{M}$ possesses a constant section curvature $\kappa \in \mathbb{R}$ if, for every $x \in \mathbb{M}$ and every $u, v \in T_x\mathbb{M}$, it holds that $K_x^{u,v} = \kappa$. In this case, the curvature endomorphism is given by

$$\mathbf{R}_x^{u,v}(w) = \kappa(u\langle v, w \rangle_x - v\langle u, w \rangle_x). \tag{224}$$

Let us illustrate the notion of constant curvature by a brief example.

Example 20. *Let us consider the hyper-sphere $\mathbb{S}^{n-1}$ endowed with its canonical metric. We already know that its curvature may be expressed as $\mathbf{R}_x^{u,v}(w) = u\langle v, w \rangle_x - v\langle u, w \rangle_x$ hence, from the property (224), we conclude immediately that the hyper-sphere is indeed a manifold of constant sectional curvature $\kappa = 1$.*

For a manifold of constant sectional curvature κ, the Jacobi equation reads

$$\nabla_v\nabla_v\xi = \kappa(\langle \xi, v \rangle v - \|v\|^2 \xi), \tag{225}$$

which is clearly a linear equation in the unknown Jacobi field ξ.

In any case, if one chooses u and v orthogonal to each other and with unit norm, it is immediately seen that $K_x^{u,v} = \langle \mathbf{R}_x^{u,v}(v), u \rangle_x$, hence the sectional curvature and the Riemannian curvature essentially coincide.

The sectional curvature plays an important role in evaluating as to whether the acceleration of the geodesic deviation will increase or decrease (locally). To formalize such observation, let us define the quantity $\ell^2 := \langle \xi, \xi \rangle$ and let us evaluate its Taylor series expansion up to the fourth order. Here ξ denotes the solution of the Jacobi equation

with initial conditions $\nabla_v \xi = w$ for $t = 0$, with $\|w\| = 1$ together with $\xi = 0$ for $t = 0$. Notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}\ell^2 = 2\langle \nabla_v \xi, \xi \rangle, \tag{226}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\ell^2 = 2\langle \nabla_v \nabla_v \xi, \xi \rangle + 2\langle \nabla_v \xi, \nabla_v \xi \rangle, \tag{227}$$

$$\frac{\mathrm{d}^3}{\mathrm{d}t^3}\ell^2 = 2\langle \nabla_v^{(3)} \xi, \xi \rangle + 6\langle \nabla_v \nabla_v \xi, \nabla_v \xi \rangle, \tag{228}$$

$$\frac{\mathrm{d}^4}{\mathrm{d}t^4}\ell^2 = 2\langle \nabla_v^{(4)} \xi, \xi \rangle + 8\langle \nabla_v^{(3)} \xi, \nabla_v \xi \rangle + 6\langle \nabla_v \nabla_v \xi, \nabla_v \nabla_v \xi \rangle. \tag{229}$$

Setting $t = 0$ gives

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\ell^2\right|_{t=0} = 2\langle \nabla_v \xi, 0 \rangle = 0, \tag{230}$$

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\ell^2\right|_{t=0} = 2\langle \nabla_v \nabla_v \xi, 0 \rangle + 2\langle w, w \rangle = 2, \tag{231}$$

$$\left.\frac{\mathrm{d}^3}{\mathrm{d}t^3}\ell^2\right|_{t=0} = 2\langle \nabla_v^{(3)} \xi, 0 \rangle + 6\langle \mathbf{R}^{v,0}(v), w \rangle = 0, \tag{232}$$

$$\left.\frac{\mathrm{d}^4}{\mathrm{d}t^4}\ell^2\right|_{t=0} = 2\langle \nabla_v^{(4)} \xi, 0 \rangle + 8\langle \nabla_v^{(3)} \xi, w \rangle + 6\langle \mathbf{R}^{v,0}(v), \mathbf{R}^{v,0}(v) \rangle = 8\langle \nabla_v^{(3)} \xi, w \rangle. \tag{233}$$

To what concerns the last term, we observe that, due to the linearity of the curvature endomorphism in its vectorial arguments, we have

$$(\nabla_v \nabla_v \nabla_v \xi)_x = (\nabla_v \mathbf{R}_x)^{v,\xi}(v) + \mathbf{R}_x^{\nabla_v v,\xi}(v) + \mathbf{R}_x^{v,\nabla_v \xi}(v) + \mathbf{R}_x^{v,\xi}(\nabla_v v). \tag{234}$$

Setting $t = 0$ and recalling that initially $\xi = 0$, we see that all terms on the right-hand side are null except for the third one, that reads $\mathbf{R}_x^{v,w}(v)$, hence

$$\left.\frac{\mathrm{d}^4}{\mathrm{d}t^4}\ell^2\right|_{t=0} = \langle \mathbf{R}_x^{v,w}(v), w \rangle_x, \tag{235}$$

therefore the Taylor expansion of the function ℓ^2 around $t = 0$ reads

$$\ell^2(t) = t^2 + \tfrac{1}{3}t^4\langle \mathbf{R}_x^{v,w}(v), w \rangle_x + \cdots, \tag{236}$$

where the dots stand for higher-order terms that we are not going to pay attention to.

Under the additional hypotheses that $\|v\| = 1$ and $\langle v, w \rangle_x = 0$, the above relationship may be rewritten in terms of the sectional curvature. In fact, by the symmetry (218) we know that $\langle \mathbf{R}_x^{v,w}(v), w \rangle_x = -\langle \mathbf{R}_x^{v,w}(w), v \rangle_x$ which, in turn, equals $-K_x^{v,w}$, therefore we may write that

$$\ell^2(t) = t^2 - \tfrac{1}{3}t^4 K_x^{v,w} + \cdots, \tag{237}$$

therefore, locally, if $K_x^{v,w} > 0$ the acceleration of geodesic deviation decreases along the fiducial geodesic, while if $K_x^{v,w} < 0$ the acceleration increases. Hence the sectional curvature determines the speed of spreading of nearby geodesics. (Therefore, in some sense, curvature acts as a force, which is the fundamental observation that motivated Einstein to formulate his General Theory of Relativity).

A further interpretation of the sectional curvature in terms of separation of nearby geodesics is worth examining. Let us define again a family of geodesics $\gamma_s(t)$ and let us assume, in particular, that these geodesic lines were shot in a peculiar way such that, if we denote again by v the tangent velocity along the fiducial geodesic and by ξ the separation vector along the same geodesic line, then it holds that $\langle v, \xi \rangle = 0$, namely the separation vector is orthogonal to the velocity vector along the fiducial geodesic (that

is to say, the geodesic curves in this family are *parallel* to one another). By appropriate parametrization, one can make sure that $\langle v, v \rangle = \langle \xi, \xi \rangle = 1$. Since $\nabla_v \nabla_v \xi$ denotes the tidal acceleration, we may interpret the effect of the curvature on the spreading of the geodesics as follows:

- whenever $\langle \nabla_v \nabla_v \xi, \xi \rangle > 0$, the tidal force is directed in the same direction of the current separation, hence nearby geodesics tend to spread even further from the fiducial one;
- whenever $\langle \nabla_v \nabla_v \xi, \xi \rangle < 0$, the tidal force is directed in the opposite direction of the current separation vector, hence nearby geodesics tend to get closer to the fiducial one;
- as opposed to the previous cases, whenever $\langle \nabla_v \nabla_v \xi, \xi \rangle = 0$, the tide does not produce any effects and the geodesics get neither closer nor farther apart (in fact, this case corresponds to zero curvature).

  Since, by the hypotheses done, $\langle \nabla_v \nabla_v \xi, \xi \rangle = \langle \mathbf{R}^{v,\xi}(v), \xi \rangle = -K^{v,\xi}$, it follows that

- in a region of *positive* sectional curvature, two nearby parallel geodesics tend to converge towards each other,
- in a region of *negative* curvature, two nearby parallel geodesics tend to deviate from one another,
- in a region of *zero* curvature, two nearby parallel geodesic curves neither converge nor diverge.

Once again, the spreading of nearby geodesics is evaluated in terms of relative acceleration, because velocity spreading is not peculiar of curvature, since even in a flat space two curves may deviate with non-zero velocity.

A neat result about sectional curvature is that the number $K_x^{u,v}$ does not actually depend on two vectors $u, v \in T_x \mathbb{M}$ but rather on the subspace spanned by such vectors, namely if we define a plane $\wp := \mathrm{span}\{u, v\}$, the sectional curvature is better written as $K_x^\wp$. This property is easily proven by defining any linear combination

$$\bar{u} = au + bv, \ \bar{v} = cu + dv, \tag{238}$$

where $a, b, c, d \in \mathbb{R}$ such that $ad \neq bc$, and calculating sectional curvature at $\bar{u}, \bar{v}$, that takes the value

$$
\begin{aligned}
K_x^{\bar{u},\bar{v}} &= \frac{\langle \mathbf{R}_x^{\bar{u},\bar{v}}(\bar{v}), \bar{u} \rangle_x}{\langle \bar{u}, \bar{u} \rangle_x \langle \bar{v}, \bar{v} \rangle_x - \langle \bar{u}, \bar{v} \rangle_x^2} \\
&= \frac{\langle \mathbf{R}_x^{au+bv,cu+dv}(cu+dv), au+bv \rangle_x}{\langle au+bv, au+bv \rangle_x \langle cu+dv, cu+dv \rangle_x - \langle au+bv, cu+dv \rangle_x^2}.
\end{aligned}
\tag{239}
$$

Given the linearity of the curvature endomorphism and of the inner product in their vectorial arguments and the symmetry properties (218), it is a little tedious but not hard to find that

$$\langle \mathbf{R}_x^{au+bv,cu+dv}(cu+dv), au+bv \rangle_x = (ad-bc)^2 \langle \mathbf{R}_x^{u,v}(v), u \rangle_x, \tag{240}$$

and that

$$
\begin{aligned}
&\langle au+bv, au+bv \rangle_x \langle cu+dv, cu+dv \rangle_x - \langle au+bv, cu+dv \rangle_x^2 \\
&= (ad-bc)^2 (\langle u, u \rangle_x \langle v, v \rangle_x - \langle u, v \rangle_x^2),
\end{aligned}
\tag{241}
$$

hence the assertion.

We conclude this section by an example that justifies the property that, in a space with constant sectional curvature, the curvature endomorphism is parallel (which is often expressed through the dry statement $\nabla \mathbf{R} = 0$).

Example 21. *Let us indicate by $u, v, w, z \in \Gamma(\mathbb{M})$ four smooth vector fields and let us form the function $\mathbf{R}_x^{u,v}(w)$. The application $x \mapsto \mathbf{R}_x^{u,v}(w)$ defines a further vector field, that may be*

covariantly-derivated along z to give $\nabla_z \mathbf{R}_x^{u,v}(w)$. *Since* $\mathbf{R}_x^{u,v}(w)$ *is non-linear in x but trilinear in its vectorial arguments, it holds that*

$$\nabla_z(\mathbf{R}_x^{u,v}(w)) = (\nabla_z \mathbf{R}_x)^{u,v}(w) + \mathbf{R}^{\nabla_z u,v}(w) + \mathbf{R}^{u,\nabla_z v}(w) + \mathbf{R}^{u,v}(\nabla_z w), \qquad (242)$$

therefore, in particular

$$(\nabla_z \mathbf{R}_x)^{u,v}(w) = \nabla_z(\mathbf{R}_x^{u,v}(w)) - \mathbf{R}^{\nabla_z u,v}(w) - \mathbf{R}^{u,\nabla_z v}(w) - \mathbf{R}^{u,v}(\nabla_z w), \qquad (243)$$

On a space with constant sectional curvature κ, *it holds that* $(\nabla_z \mathbf{R}_x)^{u,v}(w) = 0$ *for every choice of* $u, v, w, z \in \Gamma(\mathbb{M})$. *In fact, recalling that, in this case,* $\mathbf{R}_x^{u,v}(w) = \kappa(u\langle v, w\rangle_x - v\langle u, w\rangle_x)$, *we have:*

$$\begin{aligned}
(\nabla_z \mathbf{R}_x)^{u,v}(w) = {} & \kappa(\nabla_z u\langle v, w\rangle_x + u\langle \nabla_z v, w\rangle_x + u\langle v, \nabla_z w\rangle_x) \\
& - \kappa(\nabla_z v\langle u, w\rangle_x - v\langle \nabla_z u, w\rangle_x - v\langle u, \nabla_z w\rangle_x) \\
& - \kappa(\nabla_z u\langle v, w\rangle_x - v\langle \nabla_z u, w\rangle_x) \\
& - \kappa(u\langle \nabla_z v, w\rangle_x - \nabla_z v\langle u, w\rangle_x) \\
& - \kappa(u\langle v, \nabla_z w\rangle_x - v\langle u, \nabla_z w\rangle_x).
\end{aligned} \qquad (244)$$

The terms on the right-hand side cancel two by two, hence the stated property holds.

5. Continuous-Time Dynamical Systems

This section presents a survey on continuous-time dynamical systems on manifolds with particular emphasis on Lagrangian systems and non-linear damped oscillators. This survey of second-order (or double-integrator) dynamical system is presented in extrinsic terms, as per the main flow of this paper as well as, for the interested readers, in coordinates and through a special formulation based on cotangent bundles.

The present section is organized in the following way. Section 5.1 retraces calculus of variation on manifold and Section 5.2 applies such technique to derive the general equation of motion on a smooth manifold by a Lagrangian approach. The aim of Section 5.3 is to illustrate the same derivation by using the coordinate formalism, while the purpose of Section 5.4 is to retrace the same derivation by using a cotangent-bundle formalism, which is shown to be as valid as the usually-invoked tangent-bundle one. In particular, Section 5.4 digs into non-linear oscillatory phenomena by first revising classical non-linear oscillators and then showing how to extend such dynamical systems to state manifolds.

5.1. Calculus of Variations on Manifold

In the first part [13] of this multi-part tutorial, we studied a problem related to the evaluation of the first-order variation of the fundamental form $\mathscr{F} : T\mathbb{M} \to \mathbb{R}$, defined as $\mathscr{F}(x, v) := \langle v, v\rangle_x$, on a manifold and we invoked its 'extendability' to a neighborhood of the tangent bundle $T\mathbb{M}$ in the ambient bundle $\mathbb{A}^2$, namely, it is not necessary to assume the well-definedness of expressions like $\mathscr{F}(x + y, v)$ nor $\mathscr{F}(x, v + w)$. In that occasion, we mentioned that extendability provides a convenient shortcut but that it is not actually necessary to arrive at the desired result. In the present section, we aim at clarifying this point.

Variational methods constitute a core tool in applied mathematics, and with a specific weight in system theory and control, as it affords the formulation of complex dynamics through and elegant formalism and first principles. It is extremely interesting to examine variational methods on smooth manifolds from an extrinsic perspective. We shall briefly survey this topic as a preparation to the next sections were we shall use it extensively.

Let $\mathscr{F} : T\mathbb{M} \to \mathbb{R}$ denote an integrable function on a smooth manifold $\mathbb{M}$ and let $\gamma : [0, 1] \to \mathbb{M}$ denote a smooth curve on a Riemannian manifold $\mathbb{M}$. We are going to prove that the first variation of the integral (*action*) functional

$$\mathscr{A}(\gamma) := \int_0^1 \mathscr{F}(\gamma, \dot{\gamma}) \mathrm{d}t, \tag{245}$$

takes the expression:

$$\delta\mathscr{A}(\gamma) = \int_0^1 \left\langle \frac{\partial \mathscr{F}}{\partial \gamma} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{F}}{\partial \dot{\gamma}}, \delta\gamma \right\rangle^{\mathbb{A}} \mathrm{d}t, \tag{246}$$

where $\langle \star, \star \rangle^{\mathbb{A}}$ denotes the Euclidean inner product that the ambient space $\mathbb{A} \supset \mathbb{M}$ is endowed with.

To understand the present development and the properties of the symbol $\delta\gamma$ in a coordinate-free setting, let us define a smoothly-parametrized family of smooth curves $c : [-\epsilon, \epsilon] \times [0, 1] \to \mathbb{M}$. The notation $c(s, t)$ may be interpreted as follows:

- the index s determines which curve in the family is referred to; keeping the index s fixed and varying the affine parameter t, namely considering $\gamma(\star, t)$, corresponds to tracing the sth curve in the family;
- the affine parameter t determines which point on a curve is referred to; keeping t fixed and letting s vary, namely considering $\gamma(s, \star)$, corresponds to traversing the family of curves transversally in correspondence of homologous points; notice that $\gamma(s, \star)$ still traces out a curve, although not necessarily endowed with the properties of $\gamma(\star, t)$ (for example, while curves $\gamma(\star, t)$ may be geodesics, transversal curves $\gamma(s, \star)$ might not be geodesics).

Such family of smooth curves will be chosen such that

$$\begin{cases} c(0, t) = \gamma(t), \ \forall t \in [0, 1] \text{ (termed 'fiducial curve')}, \\ c(s, 0) = \gamma(0), \ c(s, 1) = \gamma(1), \ \forall s \in [-\epsilon, \epsilon], \end{cases} \tag{247}$$

namely, the family member corresponding to the index value $s = 0$ coincides to the curve γ, while the other members are free except that they share the same endpoints as γ. As mentioned, for a fixed s, the map $t \mapsto c(t, s)$ traces a smooth curve on the manifold $\mathbb{M}$. Likewise, for a fixed t, the map $s \mapsto c(t, s)$ traces a smooth 'transversal' curve on $\mathbb{M}$. Smoothness implies $\frac{\partial^2 c}{\partial t \partial s} = \frac{\partial^2 c}{\partial s \partial t}$ by the Schwarz-Clairaut theorem. Now, it pays to define the following partial derivatives:

$$\delta\gamma(t) := \left. \frac{\partial}{\partial s} c(s, t) \right|_{s=0} \text{ and } \dot{c}(s, t) := \frac{\partial}{\partial t} c(s, t). \tag{248}$$

The function $\delta\gamma$ is the tangent vector to the transversal curve at the point t, hence $\delta\gamma \in T_\gamma\mathbb{M}$. In addition, from the properties (247), it follows that $\delta\gamma(0) = \delta\gamma(1) = 0$.

On the basis of the above settings, the variation $\delta\mathscr{A}(\gamma)$ may be written as:

$$\begin{aligned} \delta \int_0^1 \mathscr{F}(\gamma, \dot{\gamma}) \mathrm{d}t &:= \lim_{h \to 0} \frac{1}{h} \left\{ \int_0^1 \mathscr{F}(c(h, t), \dot{c}(h, t)) \, \mathrm{d}t - \int_0^1 \mathscr{F}(c(0, t), \dot{c}(0, t)) \, \mathrm{d}t \right\} \\ &= \int_0^1 \lim_{h \to 0} \frac{\mathscr{F}(c(h, t), \dot{c}(h, t)) - \mathscr{F}(c(0, t), \dot{c}(0, t))}{h} \, \mathrm{d}t \\ &= \int_0^1 \left. \frac{\partial}{\partial s} \mathscr{F}(c(s, t), \dot{c}(s, t)) \right|_{s=0} \mathrm{d}t. \end{aligned} \tag{249}$$

The partial derivative within the integral in Equation (249) may be written as:

$$\left. \frac{\partial \mathscr{F}(c, \dot{c})}{\partial s} \right|_{s=0} = \left. \left\langle \frac{\partial \mathscr{F}(c, \dot{c})}{\partial c}, \frac{\partial c}{\partial s} \right\rangle^{\mathbb{A}} \right|_{s=0} + \left. \left\langle \frac{\partial \mathscr{F}(c, \dot{c})}{\partial \dot{c}}, \frac{\partial \dot{c}}{\partial s} \right\rangle^{\mathbb{A}} \right|_{s=0}. \tag{250}$$

Notice that the above expression does not require extendability to a neighbor of $T\mathbb{M}$ but merely the existence of the partial derivatives $\partial_c\mathscr{F}$ and $\partial_{\dot{c}}\mathscr{F}$. The variation of the action functional in a neighbor of the fiducial curve γ reads, therefore:

$$\delta \int_0^1 \mathscr{F}(\gamma,\dot{\gamma})\mathrm{d}t = \int_0^1 \left\langle \frac{\partial\mathscr{F}(\gamma,\dot{\gamma})}{\partial\gamma},\delta\gamma \right\rangle^{\mathbb{A}} \mathrm{d}t + \int_0^1 \left\langle \frac{\partial\mathscr{F}(\gamma,\dot{\gamma})}{\partial\dot{\gamma}},\frac{\mathrm{d}\delta\gamma}{\mathrm{d}t} \right\rangle^{\mathbb{A}} \mathrm{d}t, \qquad (251)$$

where, in the second integral, we have used the interchangeability of partial derivations permitted by the Schwartz-Clairaut theorem. In addition, integration by parts yields:

$$\int_0^1 \left\langle \frac{\partial\mathscr{F}}{\partial\dot{\gamma}},\frac{\mathrm{d}}{\mathrm{d}t}\delta\gamma \right\rangle^{\mathbb{A}} \mathrm{d}t = \left\langle \frac{\partial\mathscr{F}}{\partial\dot{\gamma}},\delta\gamma \right\rangle^{\mathbb{A}}\bigg|_0^1 - \int_0^1 \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{F}}{\partial\dot{\gamma}},\delta\gamma \right\rangle^{\mathbb{A}} \mathrm{d}t. \qquad (252)$$

The first addendum on the right-hand side equals zero since the transversal variation $\delta\gamma$ vanishes at the endpoints of the fiducial curve γ, hence the variation of the integral functional assumes the final expression (246).

A noticeable consequence of this development is that the variation $\delta\mathscr{A}(\gamma)$ depends only on the tangential components $\delta\gamma \in T_\gamma\mathbb{M}$ of the perturbation.

5.2. Second-Order Dynamical Systems on Manifold

In order to formulate second-order dynamical systems on manifolds we shall recall the Euler-Lagrange-d'Alembert modified *minimal action principle.* According to such principle, a dynamical system with which it is associated a kinetic energy function $\mathscr{K}$, a potential energy function $\mathscr{V}$ and a non-conservative force field f, generates a trajectory $\gamma : [0,\,1] \to \mathbb{M}$ on a state manifold $\mathbb{M}$ such that:

$$\delta \int_0^1 (\mathscr{K}-\mathscr{V})\mathrm{d}t + \int_0^1 \langle f,\delta\gamma\rangle_\gamma \mathrm{d}t = 0. \qquad (253)$$

Let us survey in details how the above constituent functions are defined, what is their meaning, and how to evaluate the variation of the integral. We shall assume $\mathbb{M}$ to be a smooth Riemannian manifold.

The kinetic energy $\mathscr{K} : T\mathbb{M} \to \mathbb{R}$ associated to a dynamical system is defined as

$$\mathscr{K} := \tfrac{1}{2}m\langle\dot{\gamma},\dot{\gamma}\rangle_\gamma, \qquad (254)$$

with $m > 0$. Except for the factor $\frac{m}{2}$, the kinetic energy coincides with the fundamental form $\mathscr{F}$. As such, the kinetic energy depends on the state of the system as well as on its velocity. (Notice that the 'mass' term m has been introduced to retain familiarity with well-known physics concepts. In general, in the present context it is not necessary at all because the 'mass' may be more naturally thought of as part of the metric $\langle\cdot,\cdot\rangle$, hence of the metric tensor G.)

The potential energy function associated to the dynamical system under construction is indicated by $\mathscr{V} : \mathbb{M} \to \mathbb{R}$. The potential energy is generally a function of the system's state only and, most frequently, of the Riemannian distance between the state and a reference point on the state-manifold.

The function $\mathscr{L} := \mathscr{K} - \mathscr{V}$ is referred to as *Lagrangian function* of the system. The integral of the lagrangian function along the trajectory generated by the system, namely $\mathscr{A} := \int_0^1 \mathscr{L}\,\mathrm{d}t$, is referred to as *total action* of the system. (In the international unit system, the Lagrangian is measured in Joule, while the action is measured in Joule · second.)

According to the Euler-Lagrange principle that a dynamical system is supposed to bey, in the absence of non-conservative forces acting on the system, the system generates a trajectory on the state-manifold that makes the total action stationary, namely

$$\delta \int_0^1 \mathscr{L}\,\mathrm{d}t = 0. \qquad (255)$$

There exists a further function, referred to as *Hamiltonian function* of the system, defined as

$$\mathscr{H} := \mathscr{K} + \mathscr{V}, \tag{256}$$

that represents the total energy content of the system (at any given time instant). In a system that obeys the Euler-Lagrange principle, the Hamiltonian is an invariant, namely it holds that $\frac{d}{dt}\mathscr{H} = 0$.

In the presence of non-conservative forces, the formed principle needs to be modified to accommodate the work of such forces, hence it gets extended as in (253), that we shall refer to as Euler-Lagrange-d'Alembert principle.

In order to calculate the variation of the total action, according to the method surveyed in Section 5.1, let us define a smooth family of smooth curves $\gamma : [-\epsilon, \epsilon] \times [0, 1] \to \mathbb{M}$ denotes as

$$\gamma(s,t), \ 0 \leqslant t \leqslant 1, \ -\epsilon \leqslant s \leqslant \epsilon, \tag{257}$$

where the index s individuates a member of the family, while the affine parameter t individuates a point on that curve. Let us underline that, for a fixed value of the affine parameter, the map $s \mapsto \gamma(s, \star)$ traces a smooth curve.

Let us observe that the partial derivative $\frac{\partial \gamma(s,t)}{\partial s} \in T_{\gamma(s,t)}\mathbb{M}$ represents a *transversal* tangent vector, computed as

$$\frac{\partial \gamma(s,t)}{\partial s} = \lim_{h \to 0} \frac{\gamma(s+h,t) - \gamma(s,t)}{h}. \tag{258}$$

The above-defined family of curves possesses, in particular, two fixed endpoints, namely the endpoint $t = 0$ and the endpoint $t = 1$. In formal terms, this is expressed by

$$\frac{\partial \gamma}{\partial s}(s,0) = \frac{\partial \gamma}{\partial s}(s,1) = 0, \ \forall s \in [-\epsilon, \epsilon]. \tag{259}$$

In other terms, the maps $s \mapsto \gamma(s,0)$ and $s \mapsto \gamma(s,1)$ are constants with respect to the index s. As a last definition, we set

$$\delta\gamma(t) := \frac{\partial \gamma}{\partial s}(0,t). \tag{260}$$

A consequence of the assumption (259) is that $\delta(0) = \delta(1) = 0$. The function $\delta\gamma(t)$ represents the ensemble of the variations, namely all transversal tangent vectors to all point of the fiducial curve $\gamma(t) \equiv \gamma(t,0)$. (The fiducial curve represents, in fact, the trajectory actually followed by the system.)

The variation of the total action may be defined as:

$$\delta \int_0^1 \left(\frac{1}{2}m\langle \dot\gamma, \dot\gamma \rangle_\gamma - \mathscr{V} \right) dt := \left\{ \frac{\partial}{\partial \tau} \int_0^1 \left(\frac{1}{2}m\langle \dot\gamma, G_\gamma(\dot\gamma) \rangle^\mathbb{A} - \mathscr{V} \right) dt \right\} \bigg|_{\tau=0}, \tag{261}$$

where we have used the shorthand γ to denote $\gamma(t, \tau)$. In particular, the variation of the kinetic energy may be made explicit as follows:

$$\frac{\partial}{\partial \tau} \int_0^1 \frac{1}{2}m\langle \dot\gamma, G_\gamma(\dot\gamma) \rangle^\mathbb{A} dt = \frac{1}{2}m \int_0^1 \left\langle \frac{\partial \dot\gamma}{\partial \tau}, G_\gamma(\dot\gamma) \right\rangle^\mathbb{A} dt +$$

$$\frac{1}{2}m \int_0^1 \left\langle \dot\gamma, G_\gamma^\bullet \left(\frac{\partial \gamma}{\partial \tau}, \dot\gamma \right) \right\rangle^\mathbb{A} dt + \tag{262}$$

$$\frac{1}{2}m \int_0^1 \left\langle \dot\gamma, G_\gamma \left(\frac{\partial \dot\gamma}{\partial \tau} \right) \right\rangle^\mathbb{A} dt.$$

The first and the third integrals on the right-hand side take the same value. In the second integral, it is possible to swap two tangential arguments, therefore

$$\frac{\partial}{\partial \tau} \int_0^1 \frac{1}{2} m \langle \dot\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} dt = m \int_0^1 \left\langle \frac{\partial \dot\gamma}{\partial \tau}, G_\gamma(\dot\gamma) \right\rangle^{\mathbb{A}} dt + \frac{1}{2} m \int_0^1 \left\langle \frac{\partial \gamma}{\partial \tau}, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \right\rangle^{\mathbb{A}} dt. \quad (263)$$

The first integral may be re-expressed by swapping the order of partial derivation and by applying integration by parts, which give:

$$\int_0^1 \left\langle \frac{d}{dt} \frac{\partial \gamma}{\partial \tau}, G_\gamma(\dot\gamma) \right\rangle^{\mathbb{A}} dt = \left\langle \frac{\partial \gamma}{\partial \tau}, G_\gamma(\dot\gamma) \right\rangle^{\mathbb{A}} \Big|_{t=0}^{t=1} - \int_0^1 \left\langle \frac{\partial \gamma}{\partial \tau}, \frac{d}{dt} G_\gamma(\dot\gamma) \right\rangle^{\mathbb{A}} dt, \quad (264)$$

where, in particular, we may recall from the definition of the 'bullet' derivative $G^\bullet$, that

$$\frac{d}{dt} G_\gamma(\dot\gamma) = G_\gamma^\bullet(\dot\gamma, \dot\gamma) + G_\gamma(\ddot\gamma). \quad (265)$$

Continuing with the evaluation of the variation of the integral of the kinetic energy, we get:

$$\begin{aligned}
\frac{\partial}{\partial \tau} \int_0^1 \frac{1}{2} m \langle \dot\gamma, G_\gamma(\dot\gamma) \rangle^{\mathbb{A}} dt = &- m \int_0^1 \left\langle \frac{\partial \gamma}{\partial \tau}, G_\gamma^\bullet(\dot\gamma, \dot\gamma) + G_\gamma(\ddot\gamma) \right\rangle^{\mathbb{A}} dt \\
&+ \frac{1}{2} m \int_0^1 \left\langle \frac{\partial \gamma}{\partial \tau}, G_\gamma^\bullet(\dot\gamma, \dot\gamma) \right\rangle^{\mathbb{A}} dt \\
= &\; m \int_0^1 \left\langle \frac{\partial \gamma}{\partial \tau}, -\frac{1}{2} G_\gamma^\bullet(\dot\gamma, \dot\gamma) - G_\gamma(\ddot\gamma) \right\rangle^{\mathbb{A}} dt.
\end{aligned} \quad (266)$$

Let us now evaluate the variation of the potential energy function, that reads

$$\frac{\partial}{\partial \tau} \int_0^1 \mathscr{V} dt = \int_0^1 \frac{\partial \mathscr{V}}{\partial \tau} dt = \int_0^1 \left\langle \operatorname{grad}_\gamma \mathscr{V}, \frac{\partial \gamma}{\partial \tau} \right\rangle_\gamma dt = \int_0^1 \left\langle G_\gamma(\operatorname{grad}_\gamma \mathscr{V}), \frac{\partial \gamma}{\partial \tau} \right\rangle^{\mathbb{A}} dt. \quad (267)$$

To what concerns the work of the external forcing term, ultimately it turns out that

$$\int_0^1 \langle f, \delta\gamma \rangle_\gamma dt = \int_0^1 \langle G_\gamma(f), \delta\gamma \rangle^{\mathbb{A}} dt. \quad (268)$$

Upon summing up all the above terms and setting $\tau = 0$ (wherever the variable τ is present) we obtain:

$$\begin{aligned}
\delta \int_0^1 \mathscr{L} dt + \int_0^1 \langle f, \delta\gamma \rangle_\gamma dt = &\; m \int_0^1 \left\langle -\frac{1}{2} G_\gamma^\bullet(\dot\gamma, \dot\gamma) - G_\gamma(\ddot\gamma), \delta\gamma \right\rangle^{\mathbb{A}} dt \\
&- \int_0^1 \left\langle G_\gamma(\operatorname{grad}_\gamma \mathscr{V}), \delta\gamma \right\rangle^{\mathbb{A}} dt + \int_0^1 \langle G_\gamma(f), \delta\gamma \rangle^{\mathbb{A}} dt.
\end{aligned} \quad (269)$$

Let us observe that, in the first integral on the right-hand side, the scalar product does not change if one adds any arbitrary normal term in $N_\gamma \mathbb{M}$, therefore such integral is equivalent to:

$$m \int_0^1 \left\langle -G_\gamma(\nabla_t^\gamma \dot\gamma), \delta\gamma \right\rangle^{\mathbb{A}} dt. \quad (270)$$

In conclusion, the variational principle of Euler, Lagrange and d'Alembert leads to the integral equation

$$\int_0^1 \left\langle -m\, G_\gamma(\nabla_t^\gamma \dot\gamma) - G_\gamma(\operatorname{grad}_\gamma \mathscr{V}) + G_\gamma(f), \delta\gamma \right\rangle^{\mathbb{A}} dt = 0. \quad (271)$$

Since $\delta\gamma$ is an arbitrary (tangential) perturbation, the above equation implies that the integrand must be zero. Ultimately, the calculus of the variation leads to a class of second-order dynamical systems on manifold of the type

$$m\,\nabla_t^\gamma\dot\gamma = -\mathrm{grad}_\gamma\mathscr{V} + f. \tag{272}$$

At a glance, it appears to generalize the II equation of dynamics of Newton to any smooth Riemannian manifold. In fact, let us observe that if $\mathbb{M} \equiv \mathbb{R}^p$, then $\nabla_t^\gamma\dot\gamma \equiv \ddot\gamma$, hence the above equation collapses to the familiar physics law "mass $\times$ acceleration = resultant of forces".

As an example, let us compute the energy balance of a non-conservative system (and, as a by-product, let us verify the energy conservation of a conservative system).

Example 22. *By definition, the total energy of the system is $\mathscr{H} = \mathscr{K} + \mathscr{V}$, hence the power law for the system (272) is*

$$\begin{aligned}
\frac{\mathrm{d}\mathscr{H}}{\mathrm{d}t} &= \frac{\mathrm{d}\mathscr{K}}{\mathrm{d}t} + \frac{\mathrm{d}\mathscr{V}}{\mathrm{d}t} \\
&= \tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\langle\dot\gamma,\dot\gamma\rangle_\gamma + \frac{\mathrm{d}\mathscr{V}}{\mathrm{d}t} \\
&= m\langle\nabla_t^\gamma\dot\gamma,\dot\gamma\rangle_\gamma + \langle\mathrm{grad}_\gamma\mathscr{V},\dot\gamma\rangle_\gamma \\
&= \langle m\nabla_t^\gamma\dot\gamma + \mathrm{grad}_\gamma\mathscr{V},\dot\gamma\rangle_\gamma.
\end{aligned} \tag{273}$$

From the expression (272) it follows that $m\,\nabla_t^\gamma\dot\gamma + \mathrm{grad}_\gamma\mathscr{V} = f$, and hence

$$\dot{\mathscr{H}} = \langle f,\dot\gamma\rangle_\gamma. \tag{274}$$

The dissipative components of the external force fields are those that cause a decrease of the energy content of the system, namely those focing terms such that $\langle f,\dot\gamma\rangle_\gamma < 0$ (such as, for example, the viscous damping $f := -\mu\,\dot\gamma$, with $\mu > 0$). The active components are, on the other hand, those which inject energy into the system, such as those generated by actuators and due to active damping effects.

5.3. Coordinate-Prone Lagrangian Formulation of Dynamical Systems*

As recalled, on a Riemannian manifold $\mathbb{M}$ whose tangent spaces $T_x\mathbb{M}$ are endowed with an inner product $\langle\star,\star\rangle_x$, the non-conservative Euler-Lagrange-d'Alembert principle may be stated as:

$$\int_0^1 \delta(\mathscr{K}_x(\dot x) - \mathscr{V}_x)\,\mathrm{d}t + \int_0^1 \langle f_x,\delta x\rangle_x\,\mathrm{d}t = 0, \tag{275}$$

where $f \in \Gamma(\mathbb{M})$ denotes a vector field of external forces which, in the context of mathematical system modeling and control, may include energy-dissipation terms and energy-injection terms that include control fields. On each point of the fiducial trajectory $x(t)$, the quantity δx denotes a variation with respect to the fiducial trajectory itself. A possible generalization of this principle would arise by replacing the potential function by a generalized potential [23], which would give rise to gyroscopic force fields.

On a smooth manifold $\mathbb{M}$ of dimension p, where a point $x \in \mathbb{M}$ is described by coordinates $(x^1,\ldots,x^p)$ via a coordinate chart, working with intrinsic coordinates eases the expression of the solution of the variational problem (275) in explicit form. In order to express the derivation of the general law of dynamics arising from the extended Lagrangian formulation (275), the following facts are worth summarizing:

- the standard notation for covariant and contravariant tensors indexes as well as Einstein's convention on summation indexes are in force;
- for any pair of tangent vectors $u,v \in T_x\mathbb{M}$, it holds that $\langle u,v\rangle_x = g_{ij}u^iv^j$, where g_{ij} denote the components of the metric tensor associated to the inner product $\langle\star,\star\rangle_x$,

namely, based on the canonical basis $\{\partial_1^x, \ldots, \partial_p^x\}$, $g_{ij}(x) := \langle \partial_i^x, \partial_j^x \rangle_x$; the functions g_{ij} depend on the coordinates x^k and the symmetry property $g_{ij} = g_{ji}$ holds;

- in coordinates, the Riemannian gradient of a regular function $f : \mathbb{M} \to \mathbb{R}$ at a point $x \in \mathbb{M}$ is given by:

$$(\text{grad}_x f)^h = g^{hk} \frac{\partial f}{\partial x^k};$$ (276)

the quantities $\frac{\partial f}{\partial x^k}$ denote the components of the Euclidean gradient, while the spatial functions g^{hk} denote the contravariant components of the metric tensor;

- the Christoffel symbols of the first kind associated to the Levi-Civita connection corresponding to the metric tensor of components g_{ij} are computed as:

$$\Gamma_{kji} = \frac{1}{2} \left(\frac{\partial g_{ji}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{kj}}{\partial x^i} \right);$$

the Christoffel symbols of the second kind are computed as $\Gamma_{ik}^h = g^{hj}\Gamma_{ikj}$; the Christoffel symbols of the second kind are symmetric in the covariant indices, namely, $\Gamma_{ij}^h = \Gamma_{ji}^h$; the Christoffel form $\Gamma_x : T_x\mathbb{M} \times T_x\mathbb{M} \to \mathbb{R}^r$ (a way to express the Christoffel symbol applied to two tangent directions as an array) has components $[\Gamma_x(v, w)]^k = \Gamma_{ij}^k v^i w^j$;

- the kinetic energy functional is expressed by the symmetric bilinear form $\frac{1}{2} m \langle \dot{x}, \dot{x} \rangle_x$, namely $\mathscr{K}_x(\dot{x}) = \frac{1}{2} m\, g_{ij}\dot{x}^i\dot{x}^j$, where the constant $m > 0$ plays the role of a mass term; on a Riemannian manifold, the metric tensor is positive-definite, hence $\mathscr{K}_x(\dot{x}) \geqslant 0$ on every trajectory $x(t) \in \mathbb{M}$; in the context of abstract manifold calculus, the mass term is quite immaterial since it does not represent any actual physical quantity, nevertheless, we shall leave it in the equation to help readers although, as already mentioned, it might well be incorporated into the metric tensor g;

- the potential energy function $\mathscr{V}$ depends on the coordinates x^k only;

- the variations δx^k may be chosen arbitrarily except at the boundaries of the trajectory where they vanish, namely, $\delta x^k(0) = \delta x^k(1) = 0$.

In coordinates, the principle (275) may be rewritten as:

$$\int_0^1 \delta \left(\frac{1}{2} m\, g_{ij}\dot{x}^i\dot{x}^j - \mathscr{V}_x + f_k \delta x^k \right) dt = 0,$$ (277)

where functions f_k denote the covariant components of a cotangent dissipation force field $f_x^\flat \in T_x^*\mathbb{M}$. By calculating the variation in the leftmost integral, the extended Lagrangian formulation may be expressed in the form:

$$\int_0^1 \varphi_k \delta x^k dt = 0, \quad \varphi_k := \frac{1}{2} m \frac{\partial g_{ij}}{\partial x^k} \dot{x}^i \dot{x}^j - m \frac{d}{dt}(g_{ik}\dot{x}^i) + f_k - \frac{\partial \mathscr{V}_x}{\partial x^k},$$ (278)

through integration by parts. As the variations δx^k may be chosen arbitrarily for the parameter t ranging in the open interval $(0, 1)$, the equations of dynamics on the manifold $\mathbb{M}$ are such that $\varphi_k = 0$, namely:

$$\frac{1}{2} m \frac{\partial g_{ij}}{\partial x^k} \dot{x}^i \dot{x}^j - m \frac{\partial g_{ik}}{\partial x^j} \dot{x}^i \dot{x}^j - m g_{ik}\ddot{x}^i + f_k - \frac{\partial \mathscr{V}_x}{\partial x^k} = 0.$$ (279)

Thanks to the Christoffel symbols of the first kind Γ_{ijk}, the above equation may be recast as:

$$m\, g_{ik}\ddot{x}^i + m\, \Gamma_{ijk}\dot{x}^i\dot{x}^j + \frac{\partial \mathscr{V}_x}{\partial x^k} - f_k = 0.$$ (280)

Upon multiplying both sides of the above equation by the contravariant tensor g^{hk} and upon summation with respect to the index k, the above equation takes the form:

$$m\,\ddot{x}^h + m\,g^{hk}\Gamma_{ijk}\dot{x}^i\dot{x}^j + g^{hk}\left(\frac{\partial \mathscr{V}_x}{\partial x^k} - f_k\right) = 0. \tag{281}$$

The above expression may be further condensed by noticing that:

- the terms $g^{hk}\Gamma_{ijk}$ equal the Christoffel symbols of the second kind Γ^h_{ij};
- the terms $g^{hk}\frac{\partial \mathscr{V}_x}{\partial x^k}$ denote the components of the Riemannian gradient $\mathrm{grad}_x\mathscr{V}_x$;
- the terms $g^{hk}f_k$ denote the components of the contravariant (tangent) dissipation force field $f_x \in T_x\mathbb{M}$.

In conclusion, by introducing the coordinate array $x := (x^1, x^2, \dots, x^p)$ and the velocity array $v := (v^1, v^2, \dots, v^p)$, the Equation (281) may be recast compactly as a system of first-order equations

$$\begin{cases} \dot{x} = \frac{1}{m}v, \\ \dot{v} - m\,\Gamma_x(v,v) = -\mathrm{grad}_x\mathscr{V}_x + f_x. \end{cases} \tag{282}$$

The following example treats in details the case of viscous damping as external force field and determines the rate of energy dissipation.

Example 23. *The dissipation term is often drawn from viscosity theory, namely, $f_x := -\mu\dot{x}$, with $\mu > 0$ denoting a viscous damping term. The system dynamics assumes then the expression:*

$$\begin{cases} \dot{x} = \frac{1}{m}v, \\ \dot{v} - m\,\Gamma_x(v,v) = -\mathrm{grad}_x\mathscr{V}_x - \mu\,v. \end{cases} \tag{283}$$

The dynamics (283) is governed by a driving force from a potential energy plus a viscosity term. The law of energy balance for the dissipative system (283) may be found by the equation:

$$\int_0^1 \varphi_k\dot{x}^k\mathrm{d}t = 0. \tag{284}$$

Plugging in the expression (278) gives:

$$\frac{1}{2}m\int_0^1 \frac{\partial g_{ij}}{\partial x^k}\dot{x}^k\dot{x}^i\dot{x}^j\mathrm{d}t - m\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(g_{ik}\dot{x}^i)\dot{x}^k\mathrm{d}t - \mu\int_0^1 g_{hk}\dot{x}^h\dot{x}^k\mathrm{d}t - \int_0^1 \frac{\partial \mathscr{V}_x}{\partial x^k}\dot{x}^k\mathrm{d}t = 0. \tag{285}$$

The first integral may be rewritten, again upon integration by parts, as:

$$\frac{m}{2}\int_0^1 \frac{\partial g_{ij}}{\partial x^k}\dot{x}^k\dot{x}^i\dot{x}^j\mathrm{d}t = \frac{m}{2}\int_0^1 \frac{\mathrm{d}g_{ij}}{\mathrm{d}t}\dot{x}^i\dot{x}^j\mathrm{d}t = \mathscr{K}\,|_0^1 - m\int_0^1 g_{ij}\dot{x}^i\ddot{x}^j\mathrm{d}t.$$

The second term may be written, upon integration by parts, as:

$$m\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(g_{ik}\dot{x}^i)\dot{x}^k\mathrm{d}t = 2\,\mathscr{K}\,|_0^1 - m\int_0^1 g_{ij}\dot{x}^i\ddot{x}^j\mathrm{d}t,$$

while the remaining terms may be written as

$$\mu\int_0^1 g_{ik}\dot{x}^i\dot{x}^k\mathrm{d}t + \int_0^1 \frac{\partial \mathscr{V}}{\partial x^k}\dot{x}^k\mathrm{d}t = 2\frac{\mu}{m}\int_0^1 \mathscr{K}(t)\mathrm{d}t + \mathscr{V}(t)|_0^1.$$

The energy balance equation may thus be written as:

$$\mathscr{K}(t) + \mathscr{V}(t) = \mathscr{K}(0) + \mathscr{V}(0) - 2\frac{\mu}{m}\int_0^1 \mathscr{K}(t)\mathrm{d}t, \tag{286}$$

where $\mathscr{K}(t)$ stands for $\mathscr{K}_x(\dot{x}(t))$ and $\mathscr{V}(t)$ for $\mathscr{V}_{x(t)}$. The above equation shows that the designed second-order dynamical system looses energy at a rate proportional to its kinetic energy.

A non-dissipative system ($\mu = 0$) devoid of a potential function, hence completely governed by the kinetic energy/metric, produces a trajectory that keeps the quantity that undergoes variation, namely the only kinetic energy, constant over the whole trajectory. As we already knew, therefore, the Christoffel form $\Gamma_x(\star, \star)$ may thus be computed on the basis of the variational principle:

$$\int_0^1 \delta\mathscr{K}_x \, \mathrm{d}t = 0, \tag{287}$$

invoking which is often the preferred path for computing the Christoffel form.

5.4. Cotangent Bundle Derivation of Non-Linear Oscillators on Manifolds*

Nonlinear oscillators arise from the modeling of complex physical structures [24] and make up the basis for a number of modern applications [25]. The state of nonlinear oscillators may form complex, non-repeating (although deterministic) patterns. Contributions from the scientific literature suggest how the simplest harmonic-oscillator models may be generalized to complex single or coupled oscillators [26–29].

The cotangent bundle, namely the dual $T^\star\mathbb{M}$ of the tangent bundle associated to a manifold, is as valid as the tangent bundle to formulate and study dynamical systems. We mention that some authors in manifold calculus explain the whole manifold calculus itself by starting from cotangent bundles and only later introducing tangent spaces. In the present section, we shall rather be showing how non-linear oscillators may be defined by means of tangent-bundle calculus. We shall start off by revising a number of classical oscillator models very well known in the scientific literature and shall then continue by showing how such mathematical models may be seamlessly be extended to manifolds by means of cotangent-bundle calculus.

The present section makes frequent reference to the work [30] that it is based on.

5.4.1. Classical Oscillators on Euclidean Spaces

A number of nonlinear oscillator models appear in the scientific literature as dynamical systems in a single real descriptive variable. The simplest model known in the scientific literature is perhaps the *harmonic oscillator* described by the second-order dynamical system:

$$\ddot{x} + \Omega_0^2 x = 0, \tag{288}$$

where $x \in \mathbb{R}$ denotes a descriptive variable and $\Omega_0 > 0$ denotes the natural oscillation frequency of the harmonic oscillator. The state $x = x(t)$ is a function of time, with $t \in \mathbb{R}$. As no damping is considered, the harmonic oscillator conserves its initial energy indefinitely. The same consideration applies to the simple pendulum model $\ddot{x} + \Omega_0^2 \sin x = 0$, which appears as a second-order differential equation that may be easily recast as a system of two first-order differential equations:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x}, \end{cases} \tag{289}$$

where the forcing term on the right-hand side has been written as the gradient of a potential energy function $\mathscr{V} : \mathbb{R} \to \mathbb{R}$ which, in this case, may be defined as $\mathscr{V}_x := \Omega_0^2(1 - \cos x)$. (Notice that we have therewith abandoned the notion of 'mass' which, besides of being immaterial in the present context, may always be factored out through appropriate normalization.)

A further well-studied nonlinear damped oscillator that arises from the analysis of vacuum tubes is the *van der Pol oscillator* [24], described by the second-order differential equation:

$$\ddot{x} - \mu(1 - x^2)\dot{x} + \Omega_0^2 x = 0, \tag{290}$$

where again $x \in \mathbb{R}$ denotes a descriptive variable and $\mu > 0$ denotes a damping parameter. When $\mu = 0$, the Equation (290) collapses into the equation of an harmonic oscillator (288). When $\mu > 0$, the system will eventually enter a limit cycle. Near the origin $(x, \dot{x}) = (0, 0)$ of the phase space the system is unstable, while far from the origin the system is heavily damped and hence stable. The single Equation (290) may be recast as the paired equations:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x} - \mu(x^2 - 1)v, \end{cases} \tag{291}$$

where now two variables, $x, v \in \mathbb{R}$, are used to describe its dynamics and a potential energy function $\mathscr{V} : \mathbb{R} \to \mathbb{R}$ has been introduced to quantify the forcing term, which is defined as $\mathscr{V}_x := \frac{1}{2}\Omega_0^2 x^2$, where $\Omega_0 > 0$ denotes a constant parameter. The van der Pol oscillator model may be regarded as a special case of the FitzHugh-Nagumo model [31] which, in turn, represents a simplified version of the Hodgkin-Huxley mathematical model that explains in a detailed manner the phenomena of activation and deactivation of spiking neurons. A related example arises from the examination of the *driven van der Pol oscillator*, described by the differential system:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x} - \mu(x^2 - 1)v + A\sin(\Omega t), \end{cases} \tag{292}$$

where $A > 0$ and $\Omega > 0$ denote two constant parameters. In the present case, the nonlinear damping term comprises a quantity (the driving term) that depends sinusoidally from the temporal variable t. The particular form of the damping causes a dehancement of the magnitude of strong oscillations and an amplification of the magnitude of weak oscillations. The original van der Pol oscillator has been generalized in different ways, as in the paper [32] that considers the introduction of fractional derivatives.

Another well-studied nonlinear dynamical system that exhibits a complex behavior is the *Duffing oscillator* [33]. Duffing's dynamical system reads:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x} - \mu\, v + A\sin(\Omega t), \end{cases} \tag{293}$$

where $A > 0$ and $\Omega > 0$ denote constant parameters and the potential energy function $\mathscr{V} : \mathbb{R} \to \mathbb{R}$ is defined according to one of the three known models summarized below:

- *Case of a hard Duffing oscillator*: In this case, the potential function is defined as $\mathscr{V}_x := \frac{1}{2}\Omega_0^2 x^2 + \frac{1}{4}\alpha\Omega_0^2 x^4$,
- *Case of a double-well Duffing oscillator*: In this instance, the potential is defined as $\mathscr{V}_x := -\frac{1}{2}\Omega_0^2 x^2 + \frac{1}{4}\alpha\Omega_0^2 x^4$,
- *Case of a soft Duffing oscillator*: In this case, the potential is defined as $\mathscr{V}_x := \frac{1}{2}\Omega_0^2 x^2 - \frac{1}{4}\alpha\Omega_0^2 x^4$,

where $\Omega_0 > 0$ and $\alpha > 0$ denote constants of the models. The dynamical system (293) describes the motion of a damped oscillator with a more complicated potential than in the simple harmonic oscillator. It is able to model, for example, a spring pendulum whose spring's stiffness does not exactly obey Hooke's law. A direct comparison between the Duffing oscillator (293) and the driven van der Pol oscillator (292) reveals that the linear damping term $-\mu\, v$ in the Duffing system is replaced by the non-linear damping term $-\mu(x^2 - 1)v$ in the van der Pol system. A number of generalized Duffing models have been proposed in the scientific literature. As an example, we mention the contribution [34] that

describes a generalized double-well Duffing oscillator endowed with non-linear damping described by:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\Omega_0^2 x + \alpha\Omega_0^2 x^3 - \mu\, v|v|^{\eta-1} + A\sin(\Omega t), \end{cases} \tag{294}$$

where the constant $\eta \geqslant 1$ is termed *damping exponent* and determines the degree of non-linearity of the damping term. The case $\eta = 1$ corresponds to linear damping in the system (293). A similar analysis was performed in [35] about the *universal escape oscillator* endowed with the same nonlinear damping term.

Several two-variable oscillator models are known from the literature besides the one examined in details, such as the Lotka-Volterra model of prey-predator interaction and the Wilberforce torsional pendulum. A common factor of all two-dimensional models is that their dynamics is entirely described by two variables, a positional variable x and a velocity variable v, in such a way that the phase space of these systems is formed by all pairs $(x, v) \in \mathbb{R}^2$, hence is the Euclidean space $\mathbb{R}^2$. The state space, in particular, coincides with the real line $\mathbb{R}$, and the velocity space coincides with the real line $\mathbb{R}$. This appears to be a coincidence that does not copy to the manifold case, in which the state manifold $\mathbb{M}$ and the velocity space $T\mathbb{M}$ are entirely different spaces. Clearly the tangent bundle $T\mathbb{M}$ plays a fundamental role in the description of a dynamical system on manifold, and so does its dual, namely the cotangent bundle $T^\star\mathbb{M}$.

Dynamical systems involving more than one variable are known in the scientific literature. The best known example is perhaps the *Lorenz oscillator* [36], which appears as a nonlinear, three-dimensional dynamical system that generates a complex flow. A Lorenz oscillator is described by the mathematical model:

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1), \\ \dot{x}_2 = x_1(\rho - x_3) - x_2, \\ \dot{x}_3 = x_1 x_2 - \beta x_3, \end{cases} \tag{295}$$

which includes three constant parameters $\sigma, \rho, \beta > 0$ and three variables $(x_1, x_2, x_3) \in \mathbb{R}^3$. For certain values of the three parameters, the Lorenz system exhibits complex behavior. The Lorenz equations were derived from the simplified relationships that govern convection rolls arising in the study of the atmosphere and were applied to achieve climate and weather predictions. A further noteworthy example is brought to our attention by the authors of [37], which studied the emergence of oscillations in a SIR epidemic model.

The Lorenz model (295) is of the first order, but it may be recast as a second-order model by deriving both sides of each equation with respect to the temporal parameter t:

$$\begin{cases} \ddot{x}_1 = -\sigma\dot{x}_1 + \sigma\dot{x}_2, \\ \ddot{x}_2 = (\rho - x_3)\dot{x}_1 - \dot{x}_2 - x_1\dot{x}_3, \\ \ddot{x}_3 = x_2\dot{x}_1 + x_1\dot{x}_2 - \beta\dot{x}_3. \end{cases} \tag{296}$$

By associating to the Lorenz system the state array $x := [x_1\ x_2\ x_3]^\top$, the dynamical system (296) may be cast in the compact form:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -C_x v,\ C_x := \begin{bmatrix} \sigma & -\sigma & 0 \\ x_3 - \rho & 1 & x_1 \\ -x_2 & -x_1 & \beta \end{bmatrix}. \end{cases} \tag{297}$$

The recast Lorenz system takes the form of a set of purely non-linearly damped oscillator on a flat space which may be thought of being endowed with a constant potential energy function.

A general expression for coupled scalar oscillators in $\mathbb{R}^n$ that include, as a special instance, the Lorenz model (297), may be drawn from the study [38]:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\partial_x \mathscr{V} - \varphi_x(v) + u\, f(t), \end{cases} \tag{298}$$

where $x \in \mathbb{R}^n$ denotes the state array, $\mathscr{V} : \mathbb{R}^n \to \mathbb{R}$ denotes a potential energy function, $\partial_x \mathscr{V} \in \mathbb{R}^n$ denotes the Euclidean gradient of the potential energy function, $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ indicates a linear operator in the variable v (possibly nonlinear in the variable x) that represents damping and other internal sinks/sources of energy, while the term $f : \mathbb{R} \to \mathbb{R}$ represents a forcing term in the fixed direction $u \in \mathbb{R}^n$.

5.4.2. Cotangent Bundle Notation

In the present section we shall denote the inner product of two tangent vectors through a special symbol, namely $\mathbb{G}_x : T_x\mathbb{M} \times T_x\mathbb{M} \to \mathbb{R}$. The metric 'flat' operator, denoted as $\mathbb{G}^\flat$, turns a tangent vector into a cotangent vector, namely, $\mathbb{G}_x^\flat : T_x\mathbb{M} \to T_x^\star\mathbb{M}$. Its dual is termed 'sharp' operator and is denoted by $\mathbb{G}^\sharp$. The differential of a differentiable function $\psi : \mathbb{M} \to \mathbb{R}$ at a point $x \in \mathbb{M}$ is denoted by $d_x\psi \in T_x^\star\mathbb{M}$. The Riemannian gradient of the function ψ with respect to a metric $\mathbb{G}$, evaluated at a point $x \in \mathbb{M}$ is expressed though the metric operator as $(\mathbb{G}_x^\sharp \circ d_x)\psi$. The Christoffel form associated to the metric $\mathbb{G}$ is denoted as $\overset{\mathbb{G}}{\Gamma}$. The Levi-Civita connection associated to this metric is denoted as $\overset{\mathbb{G}}{\nabla}$ and the parallel transport operator associated to the same metric is denoted as $\overset{\mathbb{G}}{\mathbb{P}}$.

Let us recall that a vector field $\varphi \in \Gamma(\mathbb{M})$ that depends on a parameter, namely $\varphi_x : T_x\mathbb{M} \to T_x\mathbb{M}$, is positive-definite if $\mathbb{G}_x(\varphi_x(v), v) > 0$ for every $v \in T_x\mathbb{M} - \{0\}$ and every $x \in \mathbb{M}$.

Let us further recall that, given two points $x, y \in \mathbb{M}$ connectable by a geodesic line $\gamma : [0, 1] \to \mathbb{M}$, their Riemannian distance is defined by

$$\mathscr{D}(x, y) := \int_0^1 \mathbb{G}_{\gamma(t)}^{\frac{1}{2}}(\dot{\gamma}(t), \dot{\gamma}(t))dt. \tag{299}$$

A fundamental result of the calculus on manifolds states that the Riemannian gradient of a squared distance function reads:

$$\mathbb{G}_x^\sharp(d_x\mathscr{D}^2(x, y)) = -2\log_x(y), \tag{300}$$

as long as the logarithm is defined. The gradient of the squared distance is known in closed form only if the expression of the inverse exponential map is known explicitly.

5.4.3. Second-Order Dynamical Systems Formulation by Cotangent Bundles

A classical dynamical system model on the ordinary space $\mathbb{R}^3$ is represented by Newton's second law of motion of a particle of mass m subjected to an external force $f : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{I} \to \mathbb{R}^3$, where $\mathbb{I}$ denotes a time interval. The trajectory $\gamma : \mathbb{I} \to \mathbb{R}^3$ followed by such particle results as the solution of Newton's equation:

$$m\frac{d^2\gamma}{dt^2} = f\left(\gamma, \frac{d\gamma}{dt}, t\right). \tag{301}$$

The second-order time-derivative $\ddot{\gamma}(t)$ quantifies the instantaneous translational acceleration of the test particle as a function of the absolute time t and the mass constant m quantifies the translational inertia of the particle. The resultant of 'external' forces f depends on the instantaneous position of the particle $\gamma(t)$, on its instantaneous translational velocity $\dot{\gamma}(t)$ and may even depend explicitly on the time.

The current section recalls a formulation of second-order dynamical systems on manifold that retraces the above-recalled dynamics of a virtual point-wise particle whose

instantaneous position is bound to belong to a Riemannian manifold $\mathbb{M}$—endowed with metric $\mathbb{G}$—rather than the ordinary space $\mathbb{R}^3$. As in the previous sections, the Hamilton's stationary-action principle, extended through the d'Alembert virtual work notion, shall be invoked to formulate the dynamics of such virtual particle. For this reason, it pays to revisit the notion notions of kinetic energy, potential energy and external excitation, mainly with the aim of clarifying the customized notation introduced at the beginning of the present section.

The *kinetic energy function* for a point-wise particle associated with the metric $\mathbb{G}$ is denoted by $\mathscr{K} : T\mathbb{M} \to \mathbb{R}$ and is defined by $\mathscr{K}_x(v) := \frac{1}{2}\mathbb{G}_x(v,v)$ for $(x,v) \in T\mathbb{M}$. On a Riemannian manifold, the metric tensor is positive-definite, hence, on every trajectory $\gamma : \mathbb{I} \to \mathbb{M}$, it holds that $\mathscr{K}_\gamma(\dot{\gamma}) \geqslant 0$. The *potential energy function* $\mathscr{V} : \mathbb{M} \to \mathbb{R}$ is assumed to depend solely on the positional variable $x \in \mathbb{M}$; whenever any exogenous forcing is lacking, the trajectory $\gamma : \mathbb{I} \to \mathbb{M}$ follows the landscape of the potential energy function. As opposed to what stated in the previous sections, in the present context an *exogenous force* at a point $x \in \mathbb{M}$ appears as a cotangent vector, namely $f_x : T_x\mathbb{M} \times \mathbb{I} \to T_x^\star\mathbb{M}$.

On the basis of the above-recalled quantities and notation, the extended stationary-action principle that determines the evolution of the state on a Riemannian manifold $\mathbb{M}$ endowed with a metric $\mathbb{G}$ takes the form:

$$\delta \int_{\mathbb{I}} (\mathscr{K}_\gamma(\dot{\gamma}) - \mathscr{V}_\gamma)\mathrm{d}t + \int_{\mathbb{I}} \mathbb{G}_\gamma(\mathbb{G}_\gamma^\sharp(f_\gamma), \delta\gamma)\mathrm{d}t = 0, \tag{302}$$

where symbol δ again denotes variation. The integrand in the leftmost integral represents the classical *Lagrangian function* of the particle and its integral represents the *total action* of the particle. The rightmost integral represents the virtual work effected by the resultant of the external driving force and represents an extension of the classical Hamilton's stationary-action principle that may be traced back to d'Alembert [39]. In the interior of the interval $\mathbb{I}$, the variation $\delta\gamma \in T_\gamma\mathbb{M}$ is arbitrary, while at the boundaries of the trajectory it vanishes to zero. Computing the variation leads to the dynamical system formulation:

$$(\mathbb{G}_\gamma^\flat \circ \overset{\mathbb{G}}{\nabla}_{\dot{\gamma}})\dot{\gamma} = -\mathrm{d}_\gamma\mathscr{V}_\gamma + f_\gamma. \tag{303}$$

By comparing the Equation (303) to Newton's law (301), it is readily seen that the term $\overset{\mathbb{G}}{\nabla}_{\dot{\gamma}}\dot{\gamma}$ represents the covariant acceleration of the particle sliding on the manifold $\mathbb{M}$, the operator $\mathbb{G}^\flat$ plays the role of inertia tensor and the term $f_x - \mathrm{d}_x\mathscr{V}_x$ is the total force that generates the motion of the particle. Of particular interest the role played by the inertia tensor $\mathbb{G}^\flat$ that replaces (and extends) the notion of inertial mass.

The covariant acceleration may be written in terms of the Christoffel form as $\overset{\mathbb{G}}{\nabla}_{\dot{\gamma}}\dot{\gamma} = \ddot{\gamma} + \overset{\mathbb{G}}{\Gamma}_\gamma(\dot{\gamma}, \dot{\gamma})$, therefore, the Equation (303) may be rewritten as the system of first-order differential equations:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\overset{\mathbb{G}}{\Gamma}_x(v, v) - (\mathbb{G}_x^\sharp \circ \mathrm{d}_x)\mathscr{V}_x + \mathbb{G}_x^\sharp(f_x). \end{cases} \tag{304}$$

in the tangent-bundle arrays $(x(t), v(t)) \in T\mathbb{M}$.

The total energy (or Hamiltonian function) $\mathscr{H} : T\mathbb{M} \times \mathbb{I} \to \mathbb{R}$ of the particle sliding on the manifold, and hence of the dynamical system (304), is defined by:

$$\mathscr{H}_x := \mathscr{K}_x + \mathscr{V}_x. \tag{305}$$

Over a trajectory $\gamma : \mathbb{I} \to \mathbb{M}$ of the system (304), it holds that:

$$\frac{\mathrm{d}\mathscr{H}_\gamma}{\mathrm{d}t} = \mathbb{G}_\gamma(\mathbb{G}_\gamma^\sharp(f_\gamma), \dot{\gamma}), \tag{306}$$

which represents a power law for the dynamical system at hand. In the case that the external forcing is absent, the system (304) is conservative as the total energy $\mathscr{H}_\gamma$ keeps constant over the trajectory, otherwise the system is non-conservative and its energy varies along its trajectory.

In the present context, the force $f_x = f_x(v, t) \in T_x^\star M$ essentially represents damping effects and external forcing terms. In particular:

- *Friction-type damping*: This kind of damping generalizes the Rayleigh damping and is expressed by the forcing term $-\frac{1}{2}\mu \frac{\partial (\mathbb{G}_x(v,v)^\epsilon)}{\partial v^\sigma} dx^\sigma$, with $\epsilon \geqslant 1$ being a damping coefficient and $\mu \geqslant 0$ being a viscosity coefficient. By computing the derivatives, it is found that the friction-type damping force equals $-\mu \|v\|_x^{2(\epsilon-1)} \mathbb{G}_x^\flat(v)$.

- *Non-linear damping*: It generalizes the nonlinear damping term that appears in the van der Pol system and assumes the expression $-(\mathbb{G}_x^\flat \circ \varphi_x)(v)$, with $\varphi \in \Gamma(\mathbb{M})$ being a vector field that depends on a parameter, such that, for each $x \in \mathbb{M}$, $\varphi_x : T_x\mathbb{M} \to T_x\mathbb{M}$ is a linear endomorphism.

- *Sinusoidal driving force*: It generalizes the notion of external sinusoidal forcing term from mono-dimensional dynamical systems. It takes the expression $A\sin(\Omega t)dx^1$, with $A, \Omega \in \mathbb{R}$ and $t \in \mathbb{I}$.

With the above assumptions, the dynamical system (304) takes the expression:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\overset{\mathbb{G}}{\Gamma}_x(v,v) - \mathbb{G}_x^\sharp(d_x\mathscr{V}_x) - \mu\|v\|_x^{2(\epsilon-1)}v - \varphi_x(v) - A\sin(\Omega t)\mathbb{G}_x^\sharp(dx^1). \end{cases} \tag{307}$$

The rate of change of the total energy due to the above forcing terms reads:

$$\frac{d\mathscr{H}_x}{dt} = -\mu\|v\|_x^{2\epsilon} - \mathbb{G}_x(\varphi_x(v), v) - A\sin(\Omega t)v^1. \tag{308}$$

If $\mu \neq 0$, the first term on the right-hand side is purely dissipative and is proportional to the kinetic energy of the system (to $\mathscr{K}_x^\epsilon$, in fact), while the second term on the right-hand side is not necessarily dissipative and may bring energy into the system. (If the vector field φ is positive-definite, then the term $-\varphi_x(v)$ in the system (307) is purely dissipative.)

Let us consider two examples that aim at clarifying the notation used and to show that the considered non-linear oscillators on manifold are indeed general enough to encompass the original classical oscillators.

Example 24. *In the special case that $\mathbb{M} = \mathbb{R}^n$, hence $T_x\mathbb{M} \equiv \mathbb{R}^n$, and $\mathbb{G}_x(v,w) = v^\top w$, it holds that $\overset{\mathbb{G}}{\Gamma}_x = 0$, $\mathbb{G}_x^\sharp(d_x\mathscr{V}_x) = \partial_x\mathscr{V}$, $\mathbb{G}_x^\sharp(dx^1) = \mathbf{1}_n$, where symbol $\mathbf{1}_n$ denotes the column array $[1, 0, \cdots, 0]^\top \in \mathbb{R}^n$. Hence, the dynamical system (307) becomes:*

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\partial_x\mathscr{V} - \mu(v^\top v)^{\epsilon-1}v - \varphi_x(v) - A\sin(\Omega t)\mathbf{1}_n, \end{cases} \tag{309}$$

where $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $\mu > 0$. The dynamical system (309) accounts for the exemplary systems (289), (291)–(294) and (298).

To further exemplify the above theoretical development, the present subsection discusses the special case of Riemannian manifolds $\mathbb{M}$ of dimension 1. The main interest on such kinds of manifolds is analytic tractability. Examples of 1-dimensional manifolds are the unit sphere $\mathbb{S}^1$ (namely, the unit circle) and the special orthogonal group $\mathbb{SO}(2)$.

In order to particularize the general Equation (304) to the case of a 1-dimensional Riemannian manifold, the following preliminary observations are in order. On a Riemannian manifold $\mathbb{M}$ of dimension 1, the tangent space $T_x\mathbb{M}$ is a linear space of dimension 1 whose basis vector is denoted by ∂_x, and the metric tensor $\mathbb{G}_x$ collapses into a scalar function $g_x > 0$, $x \in \mathbb{M}$. Given two tangent vectors $v\partial_x, w\partial_x \in T_x\mathbb{M}$, their inner product $\mathbb{G}_x(v\partial_x, w\partial_x) = g_x vw$. Consequently, the squared

norm $\|v\partial_x\|_x^2$ equals $g_x v^2$. Likewise, the cotangent space $T_x^\star \mathbb{M}$ has dimension 1 and its basis vector is denoted by $\mathrm{d}x$. Given a cotangent vector $u\,\mathrm{d}x$, it holds that $\mathbb{G}_x^\sharp(u\,\mathrm{d}x) = \frac{u}{g_x}\partial_x$. Hence, in the present example, the only Christoffel symbol $\overset{\mathbb{G}}{\Gamma}_{11}^1$ reads $\overset{\mathbb{G}}{\gamma}_x := \frac{1}{2}\frac{1}{g_x}\frac{\mathrm{d}g_x}{\mathrm{d}x} = \frac{1}{2}\frac{\mathrm{d}\log g_x}{\mathrm{d}x}$ and the only non-zero value of the Christoffel form $\overset{\mathbb{G}}{\Gamma}_x(v\partial_x, w\partial_x)$ is $\overset{\mathbb{G}}{\gamma}_x vw$. Given a potential energy function $\mathscr{V} : \mathbb{M} \to \mathbb{R}$, the only non-zero component of its gradient $\mathbb{G}_x^\sharp(\mathrm{d}_x\mathscr{V}_x)$ equals $\frac{1}{g_x}\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x}$. The general Equation (304), particularized to a 1-dimensional Riemannian manifold, read:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\overset{\mathbb{G}}{\gamma}_x v^2 - \frac{1}{g_x}\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x} - \mu(g_x v^2)^{\epsilon-1}v - \varphi_x(v). \end{cases} \tag{310}$$

An interesting special case is the one where the metric is independent of the position, namely $g_x = 1$, that implies $\overset{\mathbb{G}}{\gamma}_x = 0$. In this case, the dynamical system (310) particularizes to:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -\frac{\mathrm{d}\mathscr{V}_x}{\mathrm{d}x} - \mu v^{2(\epsilon-1)}v - \varphi_x(v). \end{cases} \tag{311}$$

The above dynamical system is a prototype for the classical nonlinear damped oscillators, such as the van der Pol oscillator and the simple pendulum.

The sphere $\mathbb{S}^1$ embedded in $\mathbb{R}^2$ may be described as the set of arrays of the form $x = [\cos\theta \ \sin\theta]^\top$, with $\theta \in \mathbb{R}$. Then, it holds that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \dot{x}\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = v\partial_x,$$

where $v := \dot{x}$ and $\partial_x := [-\sin x \ \cos x]^\top$. Choosing the Euclidean inner product on the tangent spaces $T_x\mathbb{S}^1$, it holds that $g_x = \mathbb{G}_x(\partial_x, \partial_x) = [-\sin\theta \ \cos\theta]^\top[-\sin\theta \ \cos\theta] = 1$. Hence, the dynamical equations on $\mathbb{S}^1$ endowed with the Euclidean metric are as in (311).

The exponential map corresponding to the above geometric setting may be expressed as:

$$\exp_x(v\partial_x) = \cos(v)x + \sin(v)\partial_x. \tag{312}$$

Given two points $y, z \in \mathbb{S}^1$, solving the equation $z = \exp_y(v\partial_y)$ in the unknown v gives $v = \arccos(z^\top y)$, hence:

$$\log_y(z) = \arccos(z^\top y)\partial_y, \tag{313}$$

and the geodesic distance between the points x and z is given by:

$$\mathscr{D}(z,y) = \arccos(z^\top y). \tag{314}$$

In the above expressions the symbol 'arccos' denotes the inverse cosine function. Parametrizing the point y by $[\cos\theta \ \sin\theta]^\top$ and the point z by $[\cos r \ \sin r]^\top$, the squared geodesic distance may be written as $\mathscr{D}^2(z,y) = (\theta - r)^2$.

5.4.4. Potential Energy Functions

The present subsection retraces classical potential energy functions summarized in Section 5.4.1 and extends such potentials to a general Riemannian manifold.

A well-documented potential energy function is the one arising in the study of the simple pendulum oscillating around the reference angle $r = 0$. In the classical case that $\mathbb{M} = \mathbb{R}$, it reads $\mathscr{V}_x \propto 1 - \cos x$. Such potential energy function may be extended to a general Riemannian manifold $\mathbb{M}$, endowed with a Riemannian distance function $\mathscr{D}(\star, \star)$, as:

$$\mathscr{V}_x^{(\mathrm{pen})} := \kappa(1 - \cos\mathscr{D}(x,r)), \tag{315}$$

with $\kappa > 0$ being a constant parameter and $r \in \mathbb{M}$ denoting a reference point. The potential $\mathscr{V}_x^{(\text{pen})}$ presents one of its minima in $x = r$. Rewriting the pendulum-type potential as $V_x^{(\text{pen})} = \kappa - \kappa \cos[(\mathscr{D}^2(x,r))^{\frac{1}{2}}]$ it is immediate to see that, according to the calculation rule (300), its Riemannian gradient reads:

$$\mathbb{G}_x^\sharp(\mathrm{d}_x \mathscr{V}_x^{(\text{pen})}) = -\kappa \frac{\sin \mathscr{D}(x,r)}{\mathscr{D}(x,r)} \log_x(r). \tag{316}$$

The potential energy function introduced to state the van der Pol dynamical system in the mono-dimensional state space $\mathbb{R}$ is a quadratic function centered in the reference point $r = 0$, namely $\mathscr{V}_x \propto x^2$. Such potential energy function may be extended to a general Riemannian manifold $\mathbb{M}$ by first generalizing the squared distance-to-zero x^2 by a squared distance-to-reference $(x - r)^2$ and successively by replacing the Euclidean distance by the geodesic distance, namely as

$$\mathscr{V}_x^{(\text{pol})} := \tfrac{\kappa}{2} \mathscr{D}^2(x,r), \tag{317}$$

with $\kappa > 0$ being a constant parameter and $r \in \mathbb{M}$ denoting a reference point. The potential $\mathscr{V}_x^{(\text{pol})}$ presents a single minimum in $x = r$. Moreover, according to the calculation rule (300), its Riemannian gradient reads:

$$\mathbb{G}_x^\sharp(\mathrm{d}_x \mathscr{V}_x^{(\text{pol})}) = -\kappa \log_x(r). \tag{318}$$

The contribution [40] studied an extension of the (hard) Duffing oscillator to *two* dimensions, namely, to the state space $\mathbb{R}^2$. The key point is to extend the original potential function $\mathscr{V}_x : \mathbb{R} \to \mathbb{R}$ through vector norm as:

$$\mathscr{V}_x = \tfrac{1}{2}\Omega_0^2 \|x\|^2 + \tfrac{1}{4}\kappa\Omega_0^2 \|x\|^4, \tag{319}$$

where $\| \star \|$ denotes the Euclidean norm, while $\Omega_0 > 0$ and $\kappa > 0$ are free parameters. The quantity $\|x\|$ coincides with the Euclidean distance of the state x to the reference point $r = 0$. Such observation paves the way to an extension to a general Riemannian manifold $\mathbb{M}$, that is achieved by replacing the Euclidean distance with the Riemannian distance:

$$\mathscr{V}_x^{(\text{duf})} := \pm\tfrac{1}{2}\mathscr{D}^2(x,r) \pm \tfrac{1}{4}\kappa\mathscr{D}^4(x,r), \tag{320}$$

where again the quantity $\mathscr{D}(\star, \star)$ denotes the Riemannian distance between two points on the manifold $\mathbb{M}$, the constant $\kappa > 0$ denotes a free parameter that influences the strength of the conservative forcing term corresponding to the potential, and $r \in \mathbb{M}$ represents a reference point on the state manifold around which the trajectory of the dynamical system develops. The signs $\pm$ may be chosen arbitrarily and their choice give rise to the *soft* and the *double-well* Duffing oscillator, as well as to the *hard* Duffing oscillator discussed in the paper [40]. According to the calculation rule (300), the Riemannian gradient of the extended Duffing potential $\mathscr{V}_x^{(\text{duf})}$, in the case of a hard Duffing oscillator, reads:

$$\mathbb{G}_x^\sharp(\mathrm{d}_x \mathscr{V}_x^{(\text{duf})}) = -\left(1 + \kappa\mathscr{D}^2(x,r)\right) \log_x(r). \tag{321}$$

Likewise, the *Keplerian system* mentioned in [40] may be extended to a general Riemannian manifold $\mathbb{M}$ through the introduction of the Riemannian distance $\mathscr{D}(\star, \star)$, by redefining the Keplerian potential as:

$$\mathscr{V}_x^{(\text{kep})} := -\frac{\rho}{\mathscr{D}(x,r)} + \epsilon\mathscr{D}(x,r), \tag{322}$$

with $\rho, \epsilon > 0$ denoting parameters of the potential energy function. By rewriting such Keplerian-like potential as $-\rho[\mathscr{D}^2(x,r)]^{-\frac{1}{2}} + \epsilon[\mathscr{D}^2(x,r)]^{\frac{1}{2}}$ and invoking again the calculation rule (300), it is immediate to see that its Riemannian gradient takes the expression:

$$\mathbb{G}_x^\sharp(\mathrm{d}_x\mathscr{V}_x^{(\mathrm{kep})}) = -\left(\frac{\rho}{\mathscr{D}^3(x,r)} + \frac{\epsilon}{\mathscr{D}(x,r)}\right)\log_x(r). \tag{323}$$

For exemplification purpose, let us consider two instances of non-linear damped oscillators on matrix manifolds.

Example 25. *As a first example, let us take as state space for non-linear manifold-type oscillators the manifold $\mathbb{S}^+(n)$ of symmetric positive-definite matrices, endowed with its canonical metric. If the potential energy function is written in terms of geodesic distance, its Riemannian gradient may be written by invoking the property (300). If the potential energy function is written as an explicit function of the matrix-variable $P \in \mathbb{S}^+(n)$, then the Riemannian gradient, in terms of its Gateaux derivative $\partial_P\mathscr{V}$, may be written as:*

$$\mathbb{G}_x^\sharp(\mathrm{d}_P\mathscr{V}_P) = \tfrac{1}{2}P\left(\partial_P\mathscr{V} + \partial_P^\top\mathscr{V}\right)P. \tag{324}$$

By gathering the expressions of the Christoffel operator and of the Riemannian gradient of the potential energy function, the following dynamical system equations are obtained:

$$\begin{cases} \dot{P} = U, \\ \dot{U} = UP^{-1}U - \mathbb{G}_P^\sharp(\mathrm{d}_P\mathscr{V}_P) - \mu\|U\|_P^{2(\epsilon-1)}U - A\sin(\Omega t)P\left(\mathbf{1}_{n\times n}^\top + \mathbf{1}_{n\times n}\right)P \\ \quad - \varphi_P(U). \end{cases} \tag{325}$$

In the present case, the function $\varphi_P : \mathbb{S}(n) \to \mathbb{S}(n)$.

As a second example, let us assume a Stiefel manifold is endowed with its canonical metric. In some applications, the potential energy function is written explicitly in terms of the matrix-state $X \in \mathrm{St}(n,p)$ (rather than on Riemannian distance). In the principal/minor component analysis case, for instance, the potential energy function reads $\mathscr{V}_X = \pm\tfrac{1}{2}\mathrm{tr}(X^\top\Sigma X)$, with $\Sigma \in \mathbb{S}^+(n)$, and its Gateaux derivative reads $\partial_X\mathscr{V} = \pm\Sigma X$. By the expressions of the Riemannian gradient of the potential energy function and of the Christoffel form, it is obtained the dynamical system:

$$\begin{cases} \dot{X} = U, \\ \dot{U} = -UU^\top X - XU^\top(I_n - XX^\top)U - (\partial_X\mathscr{V} - X\partial_X^\top\mathscr{V}X) - \mu\|U\|_X^{2(\epsilon-1)}U \\ \quad - \varphi_X(U) - A\sin(\Omega t)(\mathbf{1}_{n\times p} - X\mathbf{1}_{n\times p}^\top X), \end{cases} \tag{326}$$

where it must hold that $\varphi_X : T_X\mathrm{St}(n,p) \to T_X\mathrm{St}(n,p)$.

6. Control Systems on Manifolds and Numerical Implementation

This section presents a survey on error feedback control systems on manifold and their numerical implementation, with special emphasis to system synchronization.

Synchronization in a group of dynamical systems may happen forcibly or spontaneously. Spontaneous synchronization may be observed in synchronous flashing of fireflies, in flocking of fish swarms, in neuronal synchronization in central nervous system as well as in arrays of semiconductor devices. As an example of spontaneous synchronization model, we cite the Cucker-Smale system [41] that insists on the state space $\mathbb{R}^p$ and that may be expressed as

$$\ddot{x}_i = \frac{\kappa}{N}\sum_{j=1}^N \psi(\|x_i - x_j\|)(\dot{x}_j - \dot{x}_i), \tag{327}$$

where N denotes the number of individuals in the swarm, $i = 1, 2, \cdots, N$ denotes the ith individual whose state is denoted by $x_i \in \mathbb{R}^p$, $\kappa > 0$ denotes an interaction strength

constant and ψ denotes communication weight function which quantifies the degree of communications between each individual.

The classical Cucker-Smale model has been extended to manifolds, and its emergent properties have been studied, in the research paper [42]. The studied extension of the classical model (327) to a Riemannian manifold $\mathbb{M}$ reads

$$\nabla_t^{x_i} \dot{x}_i = \frac{\kappa}{N} \sum_{j=1}^{N} \psi(\mathscr{D}(x_i, x_j))(\mathbb{P}^{x_j \to x_i} \dot{x}_j - \dot{x}_i), \tag{328}$$

where the Euclidean distance in $\mathbb{R}^p$ was replaced by a Riemannian distance in $\mathbb{M}$, the state acceleration was replaced by covariant acceleration and the difference between velocities was replaced by a difference between parallel-transported velocities, which ensures that the trajectories $t \mapsto x_i(t)$ develop entirely on the state manifold $\mathbb{M}$.

An application surveyed in this section is the synchronization of the motion of two quadrotor drones (although the achieved results can be extended even to a fleet of drones). In particular, we shall survey a Lie-group control theory to allow the synchronization of the motion of two twin quadrotor drones, whose model is of second-order and insists on a Lie group. A leader drone is flying independently, while a follower should conform its dynamics in such a way that its state syncs with the state of the master. In the present context, we only focus on the rotational component of motion, which engages Lie-group control.

In order to state the quadcopter's equations of motion, since its state space is a Lie group (G, m, i, e), let us recall some relevant notation:

- A *left translation* on a Lie group $\mathbb{G}$ is denoted by $L : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ and is defined as $L_g(h) := m(i(g), h)$. Since for the quadcopter model we deal with a matrix Lie group, we may take $m(g, h) := gh$, obtaining $L_g(h) = g^{-1}h$.
- The *Lie algebra* $\mathfrak{g}$ associated to a Lie group (G, m, i, e) is a vector space endowed with *Lie brackets* $[\star, \star] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ and an *adjoint* endomorphism $\mathrm{ad}_\xi \eta := [\xi, \eta]$. The pushforward map $\mathrm{d}(L_g)_g : T_g\mathbb{G} \to \mathfrak{g}$ will be denoted as $\mathrm{d}L_g$ for brevity.
- Given a smooth function $\ell : \mathfrak{g} \to \mathbb{R}$, for a matrix Lie group we can define the *fiber derivative* of ℓ, $\frac{\delta\ell}{\delta\xi} \in \mathfrak{g}$, with $\xi \in \mathfrak{g}$, as the unique element on the algebra such that $\left\langle \frac{\delta\ell}{\delta\xi}, \eta \right\rangle_e = \mathrm{tr}((\partial_\xi \ell)^\top \eta)$ for each $\eta \in \mathfrak{g}$, where the symbol $\langle \star, \star \rangle_e$ denotes an inner product on the vector space $\mathfrak{g}$ and the symbol $\partial_\xi \ell$ denotes the Gateaux derivative (or Jacobian matrix) of the function ℓ with respect to the matrix-variable ξ.
- On a matrix Lie algebra, it holds that $\mathrm{ad}_\xi \eta := \xi\eta - \eta\xi$. In order to ease the notation in the following sections, let us define a *matrix anti-commutator* $\{\xi, \eta\} := \xi\eta + \eta\xi$, which is a symmetric form, namely $\{\xi, \eta\}^\top = \{\eta, \xi\}$, no matter what is the structure of the arguments. Conversely, the matrix commutator in the algebra is an anti-symmetric bilinear form, namely $[\xi, \eta] + [\eta, \xi] = 0$.

The present section is organized as follows. Section 6.1 presents a design and an analysis of a Lie-group synchronization theory, which is subsequently applied to synchronizing the motion of quadcopters in Section 6.2. The Section 6.3 treats an important topic, namely synchronization with vanishing control effort. To end with, Section 6.4 describes, in general terms, the notion of feedback error velocity and sheds some light between principal pushforward map and curvature endomorphism by means of homotopic nets.

6.1. Design and Analysis of a Lie-Group Synchronization Theory

The material reported in the present section is largely based on the previous contribution [43].

The Lie group considered for the quadcopter mathematical model is the special orthogonal group $\mathbb{SO}(3)$ whose Lie algebra is $\mathfrak{so}(3)$. The rotational dynamics of the leader system is governed by the equations

$$\begin{cases} \dot{R}_{\mathrm{m}} = R_{\mathrm{m}} H_{\mathrm{m}}, \\ \dot{H}_{\mathrm{m}} = \sigma(H_{\mathrm{m}}), \end{cases} \tag{329}$$

with $(R_{\mathrm{m}}, H_{\mathrm{m}}) \in \mathbb{SO}(3) \times \mathfrak{so}(3)$ denoting the leader system's state-pair, while the rotational dynamics of the controlled follower is governed by the equations

$$\begin{cases} \dot{R}_{\mathrm{s}} = R_{\mathrm{s}} H_{\mathrm{s}}, \\ \dot{H}_{\mathrm{s}} = \sigma(H_{\mathrm{s}}) + U, \end{cases} \tag{330}$$

where $(R_{\mathrm{s}}, H_{\mathrm{s}}) \in \mathbb{SO}(3) \times \mathfrak{so}(3)$ denotes the follower system's state-pair and $U \in \mathfrak{so}(3)$ denotes a control field to be designed to make the two twin systems synchronize in time. The function $\sigma : \mathfrak{so}(3) \to \mathfrak{so}(3)$ denotes the state-transition operator that ultimately describes the internal nature of the leader-follower systems twins. The mathematical model of the controlled follower system (330) will be further analyzed in Section 6.2.2 to verify the physical feasibility of the designed control strategy in relation to the effort demanded by to the actuators by the control algorithm.

The control algorithm may be designed either to synchronize the two drones' angular velocities or to synchronize their attitude. Such control goals are examined separately in what follows.

6.1.1. Velocity Synchronization

The aim of angular velocities synchronization is to ensure that

$$\lim_{t \to +\infty} ||H_{\mathrm{m}}(t) - H_{\mathrm{s}}(t)||_{\mathrm{F}} = 0. \tag{331}$$

For this purpose, a proportional-type Lie-algebra controller of the following form may be designed:

$$U := \sigma(H_{\mathrm{m}}) - \sigma(H_{\mathrm{s}}) + K_{\mathrm{P}}\, E \tag{332}$$

where $E := H_{\mathrm{m}} - H_{\mathrm{s}}$ denotes the synchronization angular-velocity mismatch and K_{P} denotes a proportional action gain coefficient. By defining $\kappa(H_{\mathrm{m}}, H_{\mathrm{s}}) := \sigma(H_{\mathrm{m}}) - \sigma(H_{\mathrm{s}}) \in \mathfrak{so}(3)$ the control field can be rewritten as

$$U := \kappa(H_{\mathrm{m}}, H_{\mathrm{s}}) + K_{\mathrm{P}}\, E. \tag{333}$$

The function $\kappa(H_{\mathrm{m}}, H_{\mathrm{s}})$ has the purpose of deleting the dynamics of the follower and to replace it with the dynamics of the leader, hence it realizes a control design principle that we refer to as *dynamics replacement*. The leader-follower pair systems will result to be velocity-synchronized as $E \approx 0$.

The convergence of the synchronization error to zero is confirmed by the following result, proven in [44]: *The control law* (332), *with* $K_{\mathrm{P}} > 0$, *synchronizes in angular velocity the follower to the leader exponentially fast.* This kind of controller does not ensure the full state synchronization for the leader-follower pair.

6.1.2. Attitude Synchronization

The problem of attitude synchronization is more involved than the problem of velocity synchronization discussed above, since a purely proportional control field would not be

effective. Attitude synchronization may be achieved by a proportional-derivative-type control. Attitude synchronization is achieved when

$$\lim_{t \to +\infty} \mathscr{D}(R_{\mathrm{m}}(t), R_{\mathrm{s}}(t)) = 0, \tag{334}$$

where $\mathscr{D} : \mathbb{SO}(3) \times \mathbb{SO}(3) \to \mathbb{R}$ denotes a Riemannian distance in $\mathbb{SO}(3)$.

　　The first step in the development of a synchronizing control algorithm is to define an *error field* that associates to each pair of leader/follower states a mismatch measure. We define an attitude synchronization error measure by the following expression

$$E := \mathrm{Log}\,(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}). \tag{335}$$

This error expression is equivalent to a 'difference' between the states R_{m} and R_{s}, written for the $\mathbb{SO}(3)$ group and belongs to the Lie algebra $\mathfrak{so}(3)$. In order to compute the time-derivative $\dot{E}$, it is convenient to make use of the following workaround:

$$\mathrm{Exp}\,(E) = R_{\mathrm{s}}^{\top} R_{\mathrm{m}} \xrightarrow{\mathrm{d}/\mathrm{d}t} \mathrm{Exp}\,(E)\dot{E} = \dot{R}_{\mathrm{s}}^{\top} R_{\mathrm{m}} + R_{\mathrm{s}}^{\top} \dot{R}_{\mathrm{m}}, \tag{336}$$

where the symbol 'Exp' denotes the matrix exponential. Replacing the expressions of the time-derivatives of the states R_{m} and R_{s} by the expressions given in (329) and (330) leads to the following relationship relating the error velocity to the attitudes and the angular velocities of the two constituent subsystems

$$\dot{E} = H_{\mathrm{m}} - (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}). \tag{337}$$

　　Deriving once more with respect to the time parameter t, and noticing that $\frac{\mathrm{d}}{\mathrm{d}t}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}) = (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}$, gives

$$\ddot{E} = \dot{H}_{\mathrm{m}} - ((R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E})^{\top} H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}) - (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} \dot{H}_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}) - (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}. \tag{338}$$

Plugging in the expressions of $\dot{H}_{\mathrm{m}}$ and $\dot{H}_{\mathrm{s}}$ from (329) and (330) yields

$$\begin{aligned} \ddot{E} = {}& \sigma(H_{\mathrm{m}}) + \dot{E}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}) - (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} (\sigma(H_{\mathrm{s}}) + U)(R_{\mathrm{s}}^{\top} R_{\mathrm{m}}) \\ & - (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}. \end{aligned} \tag{339}$$

The Lie-algebra control field U is designed so that the error dynamics obeys the law

$$\ddot{E} = -K_{\mathrm{P}} E - K_{\mathrm{D}} \dot{E}. \tag{340}$$

　　In order to express the Lie-algebra control field U in an explicit form, it is necessary to equate the expression for $\ddot{E}$ found in (339) to the desired expression of the error acceleration prescribed by the relation (340), which leads to

$$\begin{aligned} U := {}& (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\sigma(H_{\mathrm{m}})(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} - \sigma(H_{\mathrm{s}}) \\ & + (R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} H_{\mathrm{s}} - H_{\mathrm{s}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} \\ & + K_{\mathrm{P}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})E(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top} + K_{\mathrm{D}}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})\dot{E}(R_{\mathrm{s}}^{\top} R_{\mathrm{m}})^{\top}. \end{aligned} \tag{341}$$

　　The error system (340) is asymptotically convergent to a zero synchronization error. In fact, under the assumption that $K_{\mathrm{P}} > 0$ and $K_{\mathrm{D}} > 0$, the function

$$\mathscr{W} := -\frac{K_{\mathrm{P}}}{2} \mathrm{tr}(E^2) - \frac{1}{2} \mathrm{tr}(\dot{E}^2) \tag{342}$$

is Lyapunov for the error system (340). Notice that, since $\mathscr{W} = \frac{K_P}{2}||E||_F^2 + \frac{1}{2}||\dot{E}||_F^2$, from the condition $K_P > 0$, it follows that $\mathscr{W} \geqslant 0$, with equality only when $E = 0$. Upon derivation with respect to the time parameter, one gets

$$
\begin{aligned}
\dot{\mathscr{W}} &= -K_P \operatorname{tr}(E\dot{E}) - \operatorname{tr}(\dot{E}\ddot{E}) \\
&= -K_P \operatorname{tr}(E\dot{E}) - \operatorname{tr}(\dot{E}(-K_P E - K_D \dot{E})) \\
&= -K_P \operatorname{tr}(E\dot{E}) + K_P \operatorname{tr}(\dot{E}E) + K_D \operatorname{tr}(\dot{E}^2) \\
&= K_D \operatorname{tr}(\dot{E}^2).
\end{aligned}
\tag{343}
$$

Since it holds that $\dot{\mathscr{W}} = -K_D ||\dot{E}||_F^2$, from the condition $K_D > 0$, it follows that $\dot{\mathscr{W}} \leqslant 0$ with equality only when $\dot{E} = 0$.

The above result shows that the Lie-algebra control-field (341) makes the attitude of the follower subsystem (330) synchronize to that of the leader subsystem (329).

6.2. Application to Quadcopters Synchronization

The material reported in the present section is largely based on the previous contribution [45]. In the present section, we shall apply the above-recalled synchronization theory to mathematical models of quadcopters. For an introductory review, interested readers might want to consult, e.g., [46]. In particular, we are going to retrace mathematical modeling of a quadcopter drone based on the standard approach of Lagrange-Euler-d'Alembert where the exogenous generalized forces are computed on the basis of the specific structure of a quadrotor drone. The main difference between the general case considered earlier in this tutorial paper is that the Lagrange-Euler-d'Alembert principle is applied to a system whose state-space is a Lie group, which leads to the so-called Euler-Poicaré equations.

6.2.1. Lie-Group Model Formulation of a Quadrotor Drone

A quadrotor drone (also called *quadcopter*) is made of a body $\mathcal{B}$ and of four rotors $\mathcal{R}_a$, with $a = 1$, 2, 3, 4, termed *propellers*, directed upwards and placed in a square formation with equal distance from the center. Among the four propellers, rotors 2 and 4 spin clockwise, while rotors 1 and 3 spin counterclockwise [45]. Translation along a horizontal plane may be achieved by tilting the quadrotor body by modulating the rotation speed of every rotor, while a vertical displacement is achieved by modulating the total thrust exerted by the rotors on the drone's body. It is customary to attach a reference frame to the quadrotor's body, whose three orthogonal axes are denoted by x, y and z. A rotation of the body around the x axis is termed *roll*, a rotation around the y axis is termed *pitch*, while a rotation around the z axis is termed *yaw*.

The mathematical model of a quadcopter recalled here holds under a number of assumptions, namely: (1) the structure of the quadrotor is rigid and axis-symmetrical, (2) the center of the body-fixed reference-frame axes coincides with the center of mass of the quadrotor's structure, (3) the propellers are rigid, (4) the thrust and drag forces are proportional to the square of rotors' angular speed.

A mathematical model of a composite roto-translating body may be written through the well-known formalism of the Lagrange-Euler-d'Alembert principle. For a dynamical system whose state space is a Lie group, it affirms that a dynamical system generates a trajectory $[a, b] \ni t \mapsto g(t) \in \mathbb{G}$ such that

$$
\delta \int_a^b \mathscr{L}(g, \dot{g}) \, dt + \int_a^b \langle F(g, \dot{g}), \delta g \rangle_g \, dt = 0,
\tag{344}
$$

where the quantity $\mathscr{L} : T\mathbb{G} \to \mathbb{R}$ denotes a Lagrangian function for the dynamical system, the symbol $\langle \star, \star \rangle_\star$ denotes a metric for the state space, the vector field $F : T\mathbb{G} \to T\mathbb{G}$ denotes a generalized force field, the quantity $\dot{g}$ denotes the state (or configuration) velocity and the symbol δ represents a variation across two infinitely-close trajectories.

It pays now to recall the notion of *left-invariance* with special emphasis on the terms that appear in the expression (344):

- **Left-invariance of the kinetic energy**: The kinetic energy associated to a trajectory g is defined as $\mathscr{K}_g(\dot{g}) = \frac{1}{2}\langle \dot{g}, \dot{g} \rangle_g$. We shall assume the inner product representing the metric to be left-invariant, which may be expressed analytically by the condition that

$$\langle \mathrm{d}L_g(\dot{g}), \mathrm{d}L_g(\dot{g}) \rangle_e = \langle \dot{g}, \dot{g} \rangle_g, \tag{345}$$

  namely, the inner product is invariant under left-translation. In practice, this means that the inner product $\langle \dot{g}, \dot{g} \rangle_g$ does not depends explicitly on the configuration g and on the velocity $\dot{g}$, but only on their composition $\mathrm{d}L_g(\dot{g})$. Hence, the kinetic energy may be written as $\mathscr{K}_g(\dot{g}) = k(\mathrm{d}L_g(\dot{g}))$, where $k : \mathfrak{g} \to \mathbb{R}$ denotes a 'reduced' kinetic energy function.

- **Left-invariance of the potential energy**: There is no notion of left-invariance for the potential function, since it does not depend on the velocity $\dot{g}$. We shall henceforth assume (and we shall see that it is, in fact, the case) that the potential energy function for the system is simply constant with respect to the configuration g.

- **Left-invariance of the generalized force field**: In general, the product $\langle F(g, \dot{g}), \delta g \rangle_g$ depends explicitly on the quantities g and $\dot{g}$. Under the assumption that the inner product is left-invariant, we may however write that

$$\langle F(g, \dot{g}), \delta g \rangle_g = \langle \mathrm{d}_g L(F(g, \dot{g})), \mathrm{d}_g L(\delta g) \rangle_e. \tag{346}$$

To this, we add the assumption that the force field F be left-invariant, namely that $\mathrm{d}_g L(F(g, \dot{g})) = f(\mathrm{d}_g L(\dot{g}))$, where $f : \mathfrak{g} \to \mathfrak{g}$ denotes a 'reduced force field'.

In a left-invariant setting, the principle (344) can be simplified as shown in [45] to

$$\delta \int_a^b \ell(\mathrm{d}L_g(\dot{g}))\mathrm{d}t + \int_a^b \langle f(\mathrm{d}L_g(\dot{g})), \mathrm{d}L_g(\delta g) \rangle_e \mathrm{d}t = 0, \tag{347}$$

where $\ell : \mathfrak{g} \to \mathbb{R}$ represents a *reduced Lagrangian function*.

Now, the following result is instrumental in order to cast the integral principle (347) in a differential form [47]: *Let $g(s, t)$ denote a continuous family of trajectories indexed by the variable s and define $\xi := dL_g(\dot{g})$ and $\eta := dL_g(\delta g)$, where the variation δ refers to the variable s. The solution of the integral Euler-Lagrange-d'Alembert Equation (344) under perturbations of the form $\frac{\delta \xi}{\delta s} = \dot{\eta} + \mathrm{ad}_\xi \eta$ satisfies the Euler-Poincaré equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta \ell}{\delta \xi} = \mathrm{ad}_\xi^*\left(\frac{\delta \ell}{\delta \xi}\right) + f, \tag{348}$$

where the superscript * *denotes a dual operator.*

Thanks to such fundamental result, the equations that describe the motion of a dynamical system on a state Lie group $\mathbb{G}$ cast as

$$\begin{cases} \dot{g} = (dL_g)^{-1}(\xi), \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta \ell}{\delta \xi} = \mathrm{ad}_\xi^*\left(\frac{\delta \ell}{\delta \xi}\right) + f. \end{cases} \tag{349}$$

The exogenous forcing f term accounts for several contributions to the dynamics, such as energy dissipation due to friction and energy-injecting active control fields.

The Euler-Poincaré Equation (348) describes a dynamical system subjected to holonomic constraints only, subsumed by the underlying Lie group. A system subjected to non-holonomic constrains (such as a sphere rolling on a plane without slipping), may be described by a modified Euler-Poincaré equation that takes a substantially more convoluted expression [48]. The point is that the set of velocities attainable by the system is restricted

to an area of the tangent bundle termed *distribution*, which makes the equations of motion more convoluted.

The rotational component of motion of a rigid body, such as a drone, takes as descriptor an element in the Lie group $\mathbb{SO}(3)$ and in its associated Lie algebra $\mathfrak{so}(3)$. Let us now define a basis of the vector space $\mathfrak{so}(3) = \mathrm{span}(\xi_x, \xi_y, \xi_z)$ as follows:

$$\xi_x := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \ \xi_y := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \ \xi_z := \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{350}$$

Furthermore, by taking the canonical metric $\langle \xi, \eta \rangle_e := \mathrm{tr}(\xi^\top \eta)$, it is readily shown that the fiber derivative of the reduced Lagrangian reads

$$\frac{\delta \ell}{\delta \xi} = \frac{1}{2}(\partial_\xi \ell - (\partial_\xi \ell)^\top), \tag{351}$$

where $e = I_3$ denotes the identity element in the Lie group $\mathbb{SO}(3)$ and symbol ∂_x denotes, once more, a Gateaux derivative.

A quadcopter model includes two specific constants, namely its total mass and its non-standard tensor of inertia, defined as

$$M_{\mathrm{q}} := M_{\mathrm{b}} + \sum_{a=1}^{4} M_a, \ \hat{J}_{\mathrm{q}} := \hat{J}_{\mathrm{b}} + \sum_{a=1}^{4} \hat{J}_a, \tag{352}$$

where M_{b} and $\hat{J}_{\mathrm{b}}$ denote respectively the body mass of the quadcopter and its non-standard tensor of inertia, while M_a and $\hat{J}_a$, with $a = 1, 2, 3, 4$, denote respectively the mass of the rotors and their non-standard tensors of inertia. Let us recall that the non-standard tensor of inertia $\hat{J}$ of a rigid body relates to its standard tensor of inertia J by the relationship $\hat{J} = \mathrm{tr}(J)I_3 - 2J$ (we mention, with apologies, that the relation reported in [45] was incorrect of a factor $\frac{1}{2}$). It follows that the quadcopter's non-standard tensor of inertia $\hat{J}_{\mathrm{q}}$ can be deduced from the standard tensor of inertia, which is assumed to take the form of a 3×3 diagonal matrix $J_{\mathrm{q}} := \mathrm{diag}(J_x, J_y, J_z) > 0$, as follows:

$$\hat{J}_{\mathrm{q}} = \mathrm{tr}(J_{\mathrm{q}})I_3 - 2J_{\mathrm{q}} = \mathrm{diag}(-J_x + J_y + J_z, J_x - J_y + J_z, J_x + J_y - J_z). \tag{353}$$

Let us proceed on recalling the equations that describe the rotational component of motion. The angular velocities of the propellers will hereafter be denoted as $\omega_1, \omega_2, \omega_3, \omega_4$. In order to define the equations for the rotational component of motion for a quadcopter, we shall assume that all rotors are characterized by the same coefficient of rotational inertia $J_{\mathcal{R}}$ and that such rotors develop an angular velocity around the z axis $\xi_a := (-1)^a \omega_a \xi_z$, where the sign-factors $(-1)^a$ are meant to take into account that the rotors $\mathcal{R}_1$ and $\mathcal{R}_3$ spin counterclockwise, while rotors $\mathcal{R}_2$ and $\mathcal{R}_4$ spin clockwise.

Each rotor exerts a thrust, denoted as $f_a \in \mathbb{R}^3$, on the body of the quadcopter, whose magnitude is taken to be proportional to the square of the corresponding angular velocity ω_a via a lift constant $b > 0$, namely

$$f_a := \tfrac{1}{2} b \omega_a^2 \mathrm{e}_z, \ \text{where } \mathrm{e}_z := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{354}$$

(The factor $\frac{1}{2}$ has been introduced for convenience). The thrust produced by each rotor also exerts a mechanical torque on the quadcopter's body via a moment arm, denoted as $p_a \in \mathbb{R}^3$, given by

$$p_1 := \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix}, p_2 := \begin{bmatrix} 0 \\ r \\ 0 \end{bmatrix}, p_3 := \begin{bmatrix} -r \\ 0 \\ 0 \end{bmatrix}, p_4 := \begin{bmatrix} 0 \\ -r \\ 0 \end{bmatrix}, \tag{355}$$

where $r > 0$ denotes the distance between each propellers' centers and the center of mass of the quadcopter body, supposed to be the same for each propeller by symmetry.

The body of the quadcopter is subjected to further mechanical torques due to the rotation of each propeller. In fact, by virtue of the principle of angular momentum conservation, the spinning of each propeller produces a counter-torque that acts on the whole body of the quadcopter and tends to make it rotate along its z axis in the opposite direction of the resultant of the torques produced by the propellers. These torques are termed *drag torques* and their magnitude is proportional to the square of their angular velocity via a drag constant $\gamma > 0$. The resulting drag torques take the expression

$$\tau_a := \tfrac{1}{2}\gamma(-1)^a \omega_a^2 \xi_z. \tag{356}$$

(Again the factor $\frac{1}{2}$ has been introduced solely for convenience).

The resultant mechanical torque $\tau \in \mathfrak{so}(3)$ acting on the body of the quadcopter hence takes the expression

$$\tau := \overbrace{\frac{1}{2}\sum_{a=1}^{4}(-1)^a(f_a p_a^\top - p_a f_a^\top)}^{\text{Active torque}} + \overbrace{\frac{\gamma}{2}\sum_{a=1}^{4}(-1)^a \omega_a^2 \xi_z}^{\text{Drag torque}}. \tag{357}$$

By splitting such total mechanical torque with respect to the basis (350), namely as $\tau = \tau_x \xi_x + \tau_y \xi_y + \tau_z \xi_z$, one finds the components of the torques along the axes of the body-fixed reference system

$$\begin{cases} \tau_x = \tfrac{1}{2}br(\omega_4^2 - \omega_2^2), \\ \tau_y = \tfrac{1}{2}br(\omega_3^2 - \omega_1^2), \\ \tau_z = \tfrac{1}{2}\gamma(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2). \end{cases} \tag{358}$$

The components τ_x and τ_y of the resultant torque, which quantify the degree of rolling and pitching, are controlled by the unbalance of the spinning velocity of two rotors (in particular, notice that the torque along the x axis is produced by the propellers located along the y axis and vice-versa). The component τ_z of the total torque, responsible of the yawing of the quadrotor body, is produced by the unbalance between the clockwise and counterclockwise spins. Yawing is generally an unwanted effect, unless the torque τ_z is so tiny to produce a slow yawing, therefore such component is carefully kept near zero by the onboard control system.

Let us now denote by $R(t) \in \mathbb{SO}(3)$ the attitude of a quadcopter and by $\xi(t) \in \mathfrak{so}(3)$ the angular speed of this quadcopter, both thought as functions of time t. The time-evolution of the state pair $(R, \xi) \in T\mathbb{SO}(3)$ is described by the following systems of tangent-bundle differential equations:

$$\begin{cases} \dot{R} = R\xi, \\ \{\hat{J}_q, \dot{\xi}\} = [\hat{J}_q, \xi^2] + [\beta, \xi] - \dot{\beta} + 2\tau, \\ \beta := (-\omega_1 + \omega_2 - \omega_3 + \omega_4)J_\mathcal{R}\xi_z, \\ \tau := \tfrac{1}{2}br(\omega_4^2 - \omega_2^2)\xi_x + \tfrac{1}{2}br(\omega_3^2 - \omega_1^2)\xi_y + \tfrac{1}{2}\gamma(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2)\xi_z, \end{cases} \tag{359}$$

where $\{A, B\} := AB + BA$ denotes the Poisson bracket and $[A, B] := AB - BA$ denotes the Lie bracket (or simply commutator, in this case) of two matrices A, B. To confirm the above result notice that, since on the Lie algebra $\mathfrak{so}(3)$ it holds that $\mathrm{ad}_\xi^* = -\mathrm{ad}_\xi$ and on the $\mathbb{SO}(3)$ group it is verified that $dL_R^{-1}(\xi) = R\xi$, the system of Euler-Poincaré Equation (349) may be rewritten as

$$\begin{cases} \dot{R} = R\xi, \\ \frac{\mathrm{d}}{\mathrm{d}t} \frac{\delta\ell}{\delta\xi} = -\mathrm{ad}_\xi\left(\frac{\delta\ell}{\delta\xi}\right) + \tau, \end{cases} \tag{360}$$

where τ denotes the resultant of all external mechanical torques. In the present case, the reduced Lagrangian $\ell(\xi)$ of the quadrotor coincides with its kinetic energy, since there is no potential associated with the rotational component of motion. The fiber derivative of the reduced Lagrangian reads

$$\frac{\delta\ell}{\delta\xi} = \frac{1}{2}\{\hat{J}_{\mathrm{b}}, \xi\} + \frac{1}{2}\sum_{a=1}^{4}\{\hat{J}_a, \xi + \xi_a\} = \frac{1}{2}\{\hat{J}_{\mathrm{q}}, \xi\} + \frac{1}{2}\beta \tag{361}$$

where we have defined, for convenience,

$$\beta := (-\omega_1 + \omega_2 - \omega_3 + \omega_4)J_{\mathcal{R}}\xi_z. \tag{362}$$

The resultant of the external torques acting on the body of a dquadcopter drone is expressed by the relationship (358).

Whenever a drone is neither taking off nor landing, it is customary to assume the *residual rotoric angular velocity* $\Omega_{\mathrm{r}} := \omega_1 - \omega_2 + \omega_3 - \omega_4$ to be constant. Therefore, the mathematical model (359) of the rotational motion of a quadcopter may be written in the following simplified form:

$$\begin{cases} \dot{R} = R\xi, \\ \{\hat{J}_{\mathrm{q}}, \dot{\xi}\} = [\hat{J}_{\mathrm{q}}, \xi^2] + [\Omega_{\mathrm{r}}J_{\mathcal{R}}\xi_z, \xi] + 2\tau. \end{cases} \tag{363}$$

In order to describe the translational motion of a quadrotor, let us denote by $q \in \mathbb{R}^3$ the coordinate of the center of mass of the drone expressed in the inertial reference frame. The translational component of motion in a quadcopter obeys Newton's second law of dynamics written in an inertial reference frame. The linear acceleration $\ddot{q}$ is proportional to the total thrust $f_{\mathrm{T}} := \sum_{a=1}^{4} f_a$, to the gravity acceleration g and to an aerodynamic drag term. Since all the rotor thrusts are directed along the Re_z direction, the Newton law governing the translational motion can be written as:

$$M_{\mathrm{q}}\ddot{q} = \tfrac{1}{2}b(\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2)Re_z - gM_{\mathrm{q}}e_z - \Gamma\dot{q}, \tag{364}$$

where $\Gamma \in \mathbb{R}^{3\times 3}$ denotes a positive-definite (and often diagonal) aerodynamic drag tensor. The first term on the right-hand side of the above equation tells that whenever a quadcopter is horizontal to the ground (namely $R = I_3$), its center of mass can only move vertically, while whenever its body is tilted away (namely $R \neq I_3$) its center of mass can even move horizontally.

In order to write the set of equations that drive the drone's flight, it is necessary to express explicitly the Euler-Poincaré equation to cast it into the form $\dot{\xi} = \sigma(\xi, \tau)$, where $\sigma : \mathfrak{so}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3)$ denotes the state-transition function associated to this particular dynamical system. To this aim, it is useful to define a new inertia matrix

$$D := \mathrm{diag}\left(\sqrt{\frac{2J_yJ_z}{J_x}}, \sqrt{\frac{2J_xJ_z}{J_y}}, \sqrt{\frac{2J_xJ_y}{J_z}}\right). \tag{365}$$

Straightforward calculations show that $\{\hat{J}_q, \dot{\xi}\} = D\dot{\xi}D$ and that $\det(D) = 2\sqrt{2J_xJ_yJ_z} > 0$, from which it follows that the inertia matrix D is invertible. Therefore, the Euler-Poincaré equation may be rewritten explicitly as

$$\dot{\xi} = \sigma(\xi, \tau) := D^{-1}([\hat{J}_q, \xi^2] + [\beta, \xi] - \dot{\beta} + 2\tau)D^{-1}. \tag{366}$$

Ultimately, the mathematical laws that describe a quadcopter flight may be written explicitly by collecting the equations obtained in (359), (364) and (366) as

$$\begin{cases} \dot{R} = R\xi, \\ \dot{\xi} = D^{-1}([\hat{J}_q, \xi^2] + [\beta, \xi] - \dot{\beta} + 2\tau)D^{-1}, \\ \beta := (-\omega_1 + \omega_2 - \omega_3 + \omega_4)J_{\mathcal{R}}\xi_z, \\ \tau := \frac{1}{2}br(\omega_4^2 - \omega_2^2)\xi_x + \frac{1}{2}br(\omega_3^2 - \omega_1^2)\xi_y + \frac{1}{2}\gamma(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2)\xi_z, \\ \dot{q} = v, \\ \dot{v} = \frac{1}{2}\frac{b}{M_q}(\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2)Re_z - ge_z - \frac{1}{M_q}\Gamma v, \end{cases} \tag{367}$$

where the variable $v \in \mathbb{R}^3$ denotes the translational velocity of the center of mass of the quadrotor.

6.2.2. Physical Realizability of the Synchronizing Controller

The follower drone described by Equation (330) differs from the free leader drone by the application of a control field, which must be *physically realizable*, namely the control field U prescribed by the control algorithm should be effectively realizable by the propellers. A quadcopter is solely operated by changing its rotors' angular velocities, therefore it is natural to require that the values assumed by the Lie-algebra control field U must be such that it can effectively be translated into achievable angular velocities.

As a first step in a realizability analysis, it is necessary to discern, from the Equation (363) that governs the evolution of the angular velocities of the follower, those terms that represent externally-controllable torques from the terms that represent unoperable torques. Such classification follows from the physical meaning of the mathematical terms and may be written as

$$\dot{\xi} = \overbrace{D^{-1}([\hat{J}_q, \xi^2] + [\Omega_r J_{\mathcal{R}}\xi_z, \xi])D^{-1}}^{\text{Unoperable component}} + 2\,D^{-1}\tau D^{-1}, \tag{368}$$

where it is assumed that the residual rotor speed Ω_r is constant, hence the terms in β may be neglected. By defining $\sigma_{\text{uc}}(\xi) := D^{-1}([\hat{J}_q, \xi^2] + [\Omega_r J_{\mathcal{R}}\xi_z, \xi])D^{-1}$ and by equating the control field U to the operable component $2\,D^{-1}\tau D^{-1}$, the mathematical system (330) can be recast as

$$\begin{cases} \dot{R}_s = R_s\xi_s, \\ \dot{\xi}_s = \sigma_{\text{uc}}(\xi_s) + U. \end{cases} \tag{369}$$

In order to compute the angular velocities of the four propelling rotors corresponding to the control field U, it is possible to derive three conditions from the Equation (357), namely:

$$\begin{cases} br(\omega_4^2 - \omega_2^2) = \hat{\psi}_x, \\ br(\omega_3^2 - \omega_1^2) = \hat{\psi}_y, \\ \gamma(-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2) = \hat{\psi}_z, \end{cases} \tag{370}$$

where $\hat{\psi}_x$, $\hat{\psi}_y$ e $\hat{\psi}_z$ are scalar control fields defined by $\frac{1}{2}DUD =: \hat{\psi}_x\xi_x + \hat{\psi}_y\xi_y + \hat{\psi}_z\xi_z$. The system (370) is under-determined, since it totals three equations in four unknowns, therefore a fourth constraint needs to be added by considering the follower in a hovering condition. The added constraint is:

$$\frac{b}{2}(\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2) = M_q g. \tag{371}$$

The above system of equations is linear in the unknowns ω_1^2, ω_2^2, ω_3^2 and ω_4^2 and admits as solutions the angular velocities

$$
\begin{cases}
\omega_1 = \frac{1}{2\sqrt{br\gamma}}\sqrt{-2\gamma\hat{\psi}_y - br\hat{\psi}_z + 2rg\gamma M_{\mathrm{q}}}, \\
\omega_2 = \frac{1}{2\sqrt{br\gamma}}\sqrt{br\hat{\psi}_z - 2\gamma\hat{\psi}_x + 2rg\gamma M_{\mathrm{q}}}, \\
\omega_3 = \frac{1}{2\sqrt{br\gamma}}\sqrt{2\gamma\hat{\psi}_y - br\hat{\psi}_z + 2rg\gamma M_{\mathrm{q}}}, \\
\omega_4 = \frac{1}{2\sqrt{br\gamma}}\sqrt{2\gamma\hat{\psi}_x + br\hat{\psi}_z + 2rg\gamma M_{\mathrm{q}}}.
\end{cases}
\tag{372}
$$

The solutions (372) are real-valued, hence physically realizable, in principle, whenever the following two conditions are met:

- the values taken by the scalar control component $\hat{\psi}_x$, $\hat{\psi}_y$ and $\hat{\psi}_z$ do not differ too much from one another and are absolutely bounded,
- the control field components are small compared to the mass-related term $2rg\gamma M_{\mathrm{q}}$.

In practice, the spinning velocity of the propellers differs little from the spinning velocity corresponding to a hovering condition, termed *steady-state* velocity. This is defined as the spinning velocity that keeps a quadrotor still in mid-air. From the Equation (371), it is not hard to determine such propellers' spinning velocity as:

$$
\omega_{\mathrm{ss}} = \sqrt{\frac{gM_{\mathrm{q}}}{2b}} \ (\mathrm{rad/s}).
\tag{373}
$$

In order to exemplify the values of the parameters mentioned in this section, an OS4 Mini-VTOL (Vertical Take-Off and Landing) drone has been taken as reference model, as reported in [49]. The numerical values of the parameters corresponding to such quadrotor model are shown in the Table 1.

Table 1. Table recapitulating typical parameter values of a OS4 Mini-VTOL quadcopter.

Description	Parameter	Value
Overall quadrotor mass	M_{q}	650 g
Inertia on x axis	J_x	7.5×10^{-3} kg $\cdot$ m^2
Inertia along y axis	J_y	7.5×10^{-3} kg $\cdot$ m^2
Inertia along z axis	J_z	1.3×10^{-2} kg $\cdot$ m^2
Thrust coefficient	b	3.13×10^{-5} N $\cdot$ s^2
Drag coefficient	γ	7.5×10^{-7} N $\cdot$ m $\cdot$ s^2
Rotational inertia	$J_{\mathcal{R}}$	6×10^{-5} kg $\cdot$ m^2
Arm length	r	0.23 m

The value of the gravitational acceleration constant g was approximated at 9.81 m $\cdot$ s^{-2}. The calculated steady-state rotor velocity corresponding to such set of parameters is $\omega_{\mathrm{ss}} = \sqrt{gM_q/2b} \approx 319$ rad/s (roughly corresponding to 3046 revolutions per minute (RPM), quite a fast spinning). Numerical simulations show that even a fractional deviation of one RPM with respect to the steady-state velocity produces a dramatic impact on the motion of a quadrotor.

6.2.3. Numerical Recipes to Implement Synchronization of Quadrotors

We shall now illustrates a numerical technique to implement the recalled mathematical model as well as the synchronization algorithm on a computing platform.

A pseudo-code corresponding to the numerical solution of the differential equations describing the rotational component of the motion of a single drone is illustrated in the Algorithm 1. The invoked numerical integration method represents a counterpart of the well-known Euler method adapted to manifolds.

Algorithm 1 Pseudo-code to implement the (uncontrolled) mathematical model of a quadrotor on a computing platform

---

Initialize the attitude R_0 and the rotation velocity ξ_0
Set the step-size h and the number of numerical steps K to complete a simulation
for $k = 0$ to $K - 1$ **do**
    Update the attitude $R_{k+1} = R_k \, \mathrm{Exp}\,(h\,\xi_k)$
    Update the angular velocity $\xi_{k+1} = \xi_k + h\,\sigma_{\mathrm{uc}}(\xi_k)$
end for

A pseudo-code corresponding to the numerical counterpart of the differential equations describing the rotational component of the motion of a leader-follower pair of drones along with the application of a velocity-synchronization control field is illustrated in the Algorithm 2.

Algorithm 2 Pseudo-code to implement numerically a leader-follower pair and velocity synchronization by a proportional-type control algorithm

---

Initialize the attitudes $R_{m,0}$, $R_{s,0}$ and the rotation velocities $\xi_{m,0}$, $\xi_{s,0}$
Set the step-size h, the number of numerical steps K and the proportional control gain K_{P}
for $k = 0$ to $K - 1$ **do**
    Update $R_{m,k+1} = R_{m,k} \, \mathrm{Exp}\,(h\,\xi_{m,k})$
    Update $\xi_{m,k+1} = \xi_{m,k} + h\,\sigma_{\mathrm{uc}}(\xi_{m,k})$
    Evaluate the control field $U_k = \sigma_{\mathrm{uc}}(\xi_{m,k}) - \sigma_{\mathrm{uc}}(\xi_{s,k}) + K_{\mathrm{P}}(\xi_{m,k} - \xi_{s,k})$
    Update $R_{s,k+1} = R_{s,k} \, \mathrm{Exp}\,(h\,\xi_{s,k})$
    Update $\xi_{s,k+1} = \xi_{s,k} + h\,(\sigma_{\mathrm{uc}}(\xi_{s,k}) + U_k)$
end for

Moreover, a pseudo-code corresponding to the numerical implementation of attitude synchronization by a proportional-derivative-type controller is illustrated in the Algorithm 3.

Algorithm 3 Pseudo-code to implement numerically a leader-follower pair and an attitude synchronization control by a proportioanal-derivative-type controller

---

Initialize the attitudes $R_{m,0}$, $R_{s,0}$ and the rotation velocities $\xi_{m,0}$, $\xi_{s,0}$
Set the discrete step-size h, the number of numerical steps K, the proportional control gain K_{P} and the derivative control gain K_{D}
for $k = 0$ to $K - 1$ **do**
    Update $R_{m,k+1} = R_{m,k} \, \mathrm{Exp}\,(h\,\xi_{m,k})$
    Update $\xi_{m,k+1} = \xi_{m,k} + h\,\sigma_{\mathrm{uc}}(\xi_{m,k})$
    Evaluate the synchronization error $\varepsilon_k = \mathrm{Log}\,(R_{s,k}^{\top} R_{m,k})$
    Evaluate the synchronization error velocity $\dot{\varepsilon}_k = \xi_{m,k} - (R_{s,k}^{\top} R_{m,k})^{\top} \xi_{s,k} (R_{s,k}^{\top} R_{m,k})$
    Evaluate the control field (from Equation (341)) $U_k = (R_{s,k}^{\top} R_{m,k}) \sigma_{\mathrm{uc}}(\xi_{m,k}) (R_{s,k}^{\top} R_{m,k})^{\top} - \sigma_{\mathrm{uc}}(\xi_{s,k}) + (R_{s,k}^{\top} R_{m,k}) \dot{\varepsilon}_k (R_{s,k}^{\top} R_{m,k})^{\top} \xi_{s,k} - \xi_{s,k} (R_{s,k}^{\top} R_{m,k}) \dot{\varepsilon}_k (R_{s,k}^{\top} R_{m,k})^{\top} + K_{\mathrm{P}}(R_{s,k}^{\top} R_{m,k}) \varepsilon_k (R_{s,k}^{\top} R_{m,k})^{\top} + K_{\mathrm{D}}(R_{s,k}^{\top} R_{m,k}) \dot{\varepsilon}_k (R_{s,k}^{\top} R_{m,k})^{\top}$
    Update $R_{s,k+1} = R_{s,k} \, \mathrm{Exp}\,(h\,\xi_{s,k})$
    Update $\xi_{s,k+1} = \xi_{s,k} + h\,(\sigma_{\mathrm{uc}}(\xi_{s,k}) + U_k)$
end for

If the leader quadrotor is subjected to a difference in its rotor velocities that is too large (the maximum value of the difference for which the controller is effective depends on the value of the gains set), the control algorithm may loose effectiveness might hence be no longer able to synchronize the two subsystems. Such observation is in accordance with the practice: as mentioned, a quadrotor usually keeps a low difference among its propellers spinning velocities in order to be able to trace a stable trajectory.

6.3. Synchronization of Second-Order Systems on Manifolds with Vanishing Control Effort

In the previous section we surveyed synchronization of two dynamical systems on Lie groups without regard to *control effort*, which quantifies the effort demanded to the actuators by the control algorithm as well as their energy consumption. In the present section we are going to recall from [50] a control theory that ensures the control effort to vanish asymptotically as synchronization takes place.

In the present section we are studying a category of dynamical systems on a manifold $\mathbb{M}$ expressed by

$$\begin{cases} \dot{x} = v, \text{ with } x(0) = x_0 \in \mathbb{M}, \ t \geqslant 0, \\ \nabla_v v = f(x, v), \text{ with } v(0) = v_0 \in T_{x_0}\mathbb{M}, \ t \geqslant 0, \end{cases} \quad (374)$$

where $f : T\mathbb{M} \to T\mathbb{M}$ denotes a system's state-transition function and $\nabla_v v$ denotes the covariant acceleration of the state. The first equation denotes a *state-transition law*, in the manifold-valued variable x, while the second equation denotes a *velocity-transition law* in the vector field v. Since the pair (x, v) is an element of the tangent bundle, the tangent bundle $T\mathbb{M}$ represents a phase-space for the system (374).

Given a system trajectory $(t, x(t)) \in \mathbb{R} \times \mathbb{M}$ and a control vector field $(x(t), u(t)) \in T\mathbb{M}$, one may introduce a *control effort* index associated to the vector control field u as the scalar field $(x(t), \sigma(t)) \in \mathbb{M} \times \mathbb{R}_0^+$ defined by

$$\sigma := \tfrac{1}{2}\|u\|_x^2. \quad (375)$$

The scalar field σ is defined as the squared amplitude of the control field and hence quantifies the 'effort' demanded by the control algorithm to the actuators [51].

An undesirable effect of the control schemes discussed in the previous sections is that the control effort σ does not necessarily vanish after synchronization is achieved, which entails unnecessary consumption of energy (and may therefor cause premature discharge of battery). In the present section, we shall describe a control scheme, valid for a generic Riemannian manifold, capable of making the control effort vanish to zero while synchronization is being achieved. By denoting as $x_{\rm s} \in \mathbb{M}$ the state of the follower system, one such control scheme reads

$$\begin{cases} \dot{x}_{\rm s} = v_{\rm s} + u, \\ \nabla_{\dot{x}_{\rm s}} v_{\rm s} = f(x_{\rm s}, v_{\rm s}) + k\,u, \end{cases} \quad (376)$$

where $k \in \mathbb{R}$ denotes an effort-limitation gain (not to be confused with the proportional control gain) and $f : T\mathbb{M} \to T\mathbb{M}$ denotes the follower state-transition function. Notice that the follower system is described by a couple of first-order differential equations on the tangent bundle $T\mathbb{M}$ through the paired variables $(x_{\rm s}, v_{\rm s})$ and that the control field u is present not only in the equation governing the evolution of the velocity but even in the one governing the evolution of the state, as opposed to the previous synchronization scheme described by Equation (330). It is hence legitimate to wonder as to whether one such system of equations represents a physical object. To answer such question, let us recast the above two equations as a single, second-order, equation by applying the covariant differentiation operator $\nabla_{\dot{x}_{\rm s}}$ to both sides of the first equation, which gives

$$\nabla_{\dot{x}_{\rm s}} \dot{x}_{\rm s} = \nabla_{\dot{x}_{\rm s}} v_{\rm s} + \nabla_{\dot{x}_{\rm s}} u = f(x_{\rm s}, \dot{x}_{\rm s} - u) + k\,u + \nabla_{\dot{x}_{\rm s}} u. \quad (377)$$

The controlled system hence is described by the equation

$$\nabla_{\dot{x}_{\rm s}} \dot{x}_{\rm s} = f(x_{\rm s}, \dot{x}_{\rm s}) + \tilde{u}, \text{ with } \tilde{u} := k\,u + \nabla_{\dot{x}_{\rm s}} u + f(x_{\rm s}, \dot{x}_{\rm s} - u) - f(x_{\rm s}, \dot{x}_{\rm s}), \quad (378)$$

that represents a physical (second-order or double-integrator) dynamical system controlled by an effective control field $\tilde{u}$ that appears to depend on a primary control field u. As an

instance, if the primary control u is proportional to an error feedback and the state-transition function is independent of the velocity field v_{s}, then the effective control field $\tilde{u}$ is of proportional-derivative type.

It is worth emphasizing that the constant k plays a fundamental role in ensuring the asymptotic vanishing to zero of the control efforts. Before getting to the main result of this section, it is surely instructive to take a look at a couple of simpler cases.

Example 26. *Assume that $\mathbb{M} \equiv \mathbb{R}^n$ and that the function f is continuous in the first argument and Lipschitz-continuous in the second argument, with Lipschitz constant L_f. In this case, we shall consider the control system*

$$
\begin{cases}
\dot{x}_{\mathrm{m}} = v_{\mathrm{m}} \\
\dot{v}_{\mathrm{m}} = f(x_{\mathrm{m}}, v_{\mathrm{m}}) \\
\dot{x}_{\mathrm{s}} = v_{\mathrm{s}} + u, \\
\dot{v}_{\mathrm{s}} = f(x_{\mathrm{s}}, v_{\mathrm{s}}) + k\,u, \\
u := v_{\mathrm{m}} - v_{\mathrm{s}} - \frac{c}{2}(x_{\mathrm{s}} - x_{\mathrm{m}}), \quad \textbf{(Primary control field)}
\end{cases}
\quad
\begin{array}{l}
\textbf{(Free leader)} \\[1.2em]
\textbf{(Controlled follower)}
\end{array}
\tag{379}
$$

where $c > 0$ is a constant (that determines the synchronization speed). The control field 'u' makes the follower synchronize to the leader exponentially fast. Moreover, a main finding is that, provided $k \neq L_f + \frac{c}{4}$ and $k > L_f$, the control effort asymptotically vanishes to zero.

In order to show that the above control scheme is synchronizing, let us adopt again a Lyapunov-function perspective and define

$$
\begin{cases}
\mathscr{W} := \frac{1}{2}(x_{\mathrm{m}} - x_{\mathrm{s}})^{\top}(x_{\mathrm{m}} - x_{\mathrm{s}}), \\
\sigma := \frac{1}{2} u^{\top} u.
\end{cases}
\tag{380}
$$

It is not hard to show that $\mathscr{W}$ is indeed a Lyapunov function for the controller-follower subsystem, in fact:

$$
\begin{aligned}
\dot{\mathscr{W}} &= (x_{\mathrm{m}} - x_{\mathrm{s}})^{\top}(\dot{x}_{\mathrm{m}} - \dot{x}_{\mathrm{s}}) \\
&= (x_{\mathrm{m}} - x_{\mathrm{s}})^{\top}(v_{\mathrm{m}} - v_{\mathrm{s}} - u) \\
&= (x_{\mathrm{m}} - x_{\mathrm{s}})^{\top}\left(v_{\mathrm{m}} - v_{\mathrm{s}} - v_{\mathrm{m}} + v_{\mathrm{s}} + \frac{c}{2}(x_{\mathrm{s}} - x_{\mathrm{m}})\right) \\
&= -\frac{c}{2}(x_{\mathrm{m}} - x_{\mathrm{s}})^{\top}(x_{\mathrm{m}} - x_{\mathrm{s}}) \\
&= -c\,\mathscr{W}.
\end{aligned}
\tag{381}
$$

The last equation implies that $\|x_{\mathrm{m}}(t) - x_{\mathrm{s}}(t)\| = \mathscr{D}_0\, e^{-ct/2}$, where $\|\cdot\|$ denotes the Euclidean vector norm and $\mathscr{D}_0 := \|x_{\mathrm{m}}(0) - x_{\mathrm{s}}(0)\|$. Therefore, the inter-state distance tends asymptotically to zero with an exponential speed (namely, quickly at the beginning and slowly while approaching synchronization).

In order to ascertain the dynamics of the control effort, it pays to rewrite the expression of the control field in two equally meaningful ways, namely

$$
v_{\mathrm{m}} - v_{\mathrm{s}} = u + \frac{c}{2}(x_{\mathrm{s}} - x_{\mathrm{m}}),
\tag{382}
$$

$$
v_{\mathrm{s}} + u - v_{\mathrm{m}} = -\frac{c}{2}(x_{\mathrm{s}} - x_{\mathrm{m}}).
\tag{383}
$$

The above two equations imply, respectively, that

$$
\|v_{\mathrm{m}} - v_{\mathrm{s}}\| \leqslant \|u\| + \frac{c}{2}\|x_{\mathrm{s}} - x_{\mathrm{m}}\|,
\tag{384}
$$

$$
\dot{x}_{\mathrm{s}} - \dot{x}_{\mathrm{m}} = -\frac{c}{2}(x_{\mathrm{s}} - x_{\mathrm{m}}).
\tag{385}
$$

The dynamics of the control effort is hence found to be governed by the following relation:

$$
\begin{aligned}
\dot{\sigma} &= u^\top \dot{u} \\
&= u^\top \left(\dot{v}_{\mathrm{m}} - \dot{v}_{\mathrm{s}} - \tfrac{c}{2}(\dot{x}_{\mathrm{s}} - \dot{x}_{\mathrm{m}}) \right) \\
&= u^\top \left(f(x_{\mathrm{m}}, v_{\mathrm{m}}) - f(x_{\mathrm{s}}, v_{\mathrm{s}}) - k\,u - \tfrac{c}{2}(\dot{x}_{\mathrm{s}} - \dot{x}_{\mathrm{m}}) \right) \\
&= -k\,u^\top u + u^\top (f(x_{\mathrm{m}}, v_{\mathrm{m}}) - f(x_{\mathrm{s}}, v_{\mathrm{s}})) + \tfrac{1}{4}c^2 u^\top (x_{\mathrm{s}} - x_{\mathrm{m}}),
\end{aligned}
\tag{386}
$$

with initial condition $\sigma(0) = \sigma_0$ and with $t \geqslant 0$. Since the function f was assumed to be continuous with respect to its x-argument and Lipschitz-continuous with respect to its v-argument, for all constants $\varepsilon > 0$, vectors v_{m} and v_{s}, there exist two constants $\delta, L_f > 0$ such that, for every $\dot{x}_{\mathrm{m}}, \dot{x}_{\mathrm{s}}$ such that $\|\dot{x}_{\mathrm{m}} - \dot{x}_{\mathrm{s}}\| < \delta$, it results that $\|f(x_{\mathrm{m}}, v_{\mathrm{m}}) - f(x_{\mathrm{s}}, v_{\mathrm{s}})\| < \varepsilon + L_f \|v_{\mathrm{m}} - v_{\mathrm{s}}\|$. By taking profit of such inequality and of the Cauchy-Schwarz inequality about the inner product of two vectors, the right-hand terms in the differential Equation (386) may be approximated to give the following differential inequation:

$$
\dot{\sigma} < -2k\sigma + \sqrt{2\sigma}(\varepsilon + L_f \|v_{\mathrm{m}} - v_{\mathrm{s}}\|) + \tfrac{c^2}{4}\sqrt{2\sigma}\|x_{\mathrm{s}} - x_{\mathrm{m}}\|.
\tag{387}
$$

By invoking the relationships (384) and (385), from the above inequality one obtains

$$
\begin{aligned}
\dot{\sigma} &< -2k\sigma + \sqrt{2\sigma}\left(\varepsilon + L_f\|u\| + \tfrac{1}{2}cL_f\|x_{\mathrm{s}} - x_{\mathrm{m}}\| \right) + \tfrac{1}{4}c^2\sqrt{2\sigma}\|x_{\mathrm{s}} - x_{\mathrm{m}}\| \\
&= (-2k + 2L_f)\sigma + \left(\varepsilon + \tfrac{1}{2}cL_f \mathscr{D}_0 e^{-ct/2} + \tfrac{1}{4}c^2 \mathscr{D}_0 e^{-ct/2} \right) \sqrt{2\sigma}.
\end{aligned}
\tag{388}
$$

For the sake of notation conciseness, it is convenient to define the following two quantities

$$
\begin{cases}
k^* := 2(k - L_f), \\
\rho(t) := \varepsilon + \left(\tfrac{cL_f}{2} + \tfrac{c^2}{4} \right) \mathscr{D}_0 e^{-ct/2},
\end{cases}
\tag{389}
$$

in such a way that the above differential inequation may be rewritten compactly as

$$
\dot{\sigma} + k^*\sigma < \rho\sqrt{2\sigma}.
\tag{390}
$$

In view of applying a Grönwall argument [52], it is convenient to study the associated differential equation which, upon the change of variable $2\sigma =: \xi^2$, may be cast as

$$
\dot{\xi}(t) + k^*\xi(t) = \rho(t).
\tag{391}
$$

This is a first-order, inhomogeneous linear differential equation with initial condition $\xi(0) = \sqrt{2\sigma_0}$ whose closed-form solution reads

$$
\xi(t) = \xi_* e^{-k^* t} + \frac{\mathscr{D}_0}{2k^* - c}\left(cL_f + \frac{c^2}{2} \right) e^{-ct/2} + \frac{\varepsilon}{k^*},
\tag{392}
$$

where ξ_* depends from the initial condition $\xi(0)$ and is given by

$$
\xi_* := \xi_0 - \frac{\varepsilon}{k^*} - \frac{c\,\mathscr{D}_0}{2k^* - c}\left(L_f + \frac{c}{2} \right).
\tag{393}
$$

Note that, by hypothesis, $k^\star > 0$ and $k^\star \neq \tfrac{c}{2}$, hence the above quantities are well-written. Therefore, by the Grönwall's inequality the solution of the differential inequation (390) satisfies

$$
\sigma < \frac{1}{2}\left\{ \xi_* e^{-k^* t} + \frac{c\,\mathscr{D}_0}{2k^* - c}\left(L_f + \frac{c}{2} \right) e^{-ct/2} + \frac{\varepsilon}{k^*} \right\}^2.
\tag{394}
$$

The first and the second term within the square brackets converge to zero. Since the quantity ε is arbitrary, it may be assumed as small as desired, from which the assertion follows.

Since the two exponential functions within the square brackets peter out with different speed, the vanishing speed of the control effort to zero is dominated by the value $\max\left\{k^\star, \frac{c}{2}\right\}$.

The second simpler result examines the case that $\mathbb{M}$ is a general manifold but under the assumption that the leader subsystem and the follower subsystem are perfectly synchronized.

Example 27. *Consider the coupled dynamical systems*

$$
\begin{cases}
\dot{x}_\mathrm{m} = v_\mathrm{m}, \\
\nabla_{\dot{x}_\mathrm{m}} v_\mathrm{m} = f(x_\mathrm{m}, v_\mathrm{m}), \\
\dot{x}_\mathrm{s} = v_\mathrm{s} + u, \\
\nabla_{\dot{x}_\mathrm{s}} v_\mathrm{s} = f(x_\mathrm{s}, v_\mathrm{s}) + k\, u, \\
u := \mathbb{P}^{x_\mathrm{m} \to x_\mathrm{s}}(v_\mathrm{m} - \frac{c}{2}\log_{x_\mathrm{m}} x_\mathrm{s}) - v_\mathrm{s}
\end{cases}
\tag{395}
$$

evolving on a Riemannian manifold $\mathbb{M}$ and assume that the velocity-transition function f is Lipschitz continuous in the second argument, uniformly in the first argument, with Lipschitz constant L_f. The state manifold is supposed to be equipped with a Levi-Civita connection ∇. At perfect synchronization, the control scheme (395) with $k > L_f$ makes the control effort vanish to zero at least exponentially fast.

In order to justify such statement, assume that the leader subsystem and the follower subsystem are perfectly synchronized, namely, that $x_\mathrm{s} = x_\mathrm{m} := x$, $\dot{x}_\mathrm{s} = \dot{x}_\mathrm{m} := \dot{x}$. Notice that, even at perfect synchronization, one must assume that $v_\mathrm{s} \neq v_\mathrm{m}$, in general, although the pair $v_\mathrm{s}, v_\mathrm{m}$ belongs to the same tangent space $T_x\mathbb{M}$. At perfect synchronization, the control vector field $u = \mathbb{P}^{x_\mathrm{m} \to x_\mathrm{s}}\left(v_\mathrm{m} - \frac{c}{2}\log_{x_\mathrm{m}}(x_\mathrm{s})\right) - v_\mathrm{s}$ simplifies to $u = v_\mathrm{m} - v_\mathrm{s} \in T_x\mathbb{M}$, therefore it holds that

$$
\dot{\sigma} = \langle u, \nabla_{\dot{x}} u \rangle_x = \langle u, \nabla_{\dot{x}} v_\mathrm{m} - \nabla_{\dot{x}} v_\mathrm{s} \rangle_x = \langle u, f(x, v_\mathrm{m}) - f(x, v_\mathrm{s}) - k\, u \rangle_x.
\tag{396}
$$

By the hypothesis made, the function f satisfies the Lipschitz inequality in the space $T_x\mathbb{M}$

$$
\|f(x, v_\mathrm{m}) - f(x, v_\mathrm{s})\|_x \leqslant L_f \|v_\mathrm{m} - v_\mathrm{s}\|_x = L_f \|u\|_x,
\tag{397}
$$

with the constant L_f being independent from the state x; therefore, from the relation (396), by applying the Cauchy-Schwarz inequality for the inner product and by recalling the above Lipschitz continuity condition, we may obtain the following differential inequation:

$$
\dot{\sigma} \leqslant L_f \|u\|_x^2 - k\langle u, u \rangle_x = 2(L_f - k)\sigma,
\tag{398}
$$

with initial condition $\sigma(0) = \sigma_0$ and where it is understood that $t \geqslant 0$. Set $k^\star := 2(k - L_f)$ as effective decay constant. By the Grönwall's inequality, the above relation implies that the actual control effort magnitude $\sigma(t)$ is always bounded from above by the exponential curve $\sigma_0 e^{-k^\star t}$, hence the control effort vanishes to zero at least exponentially fast.

In the equations that describe second-order dynamical systems discussed in the paper [30], the velocity-transition function appears as the summation of a gradient term (that depends only on the system's state) and of a number of non-linear damping terms that depend on the system's state velocity and on the system's state itself. We should underline that the assumption that the velocity-transition function appears Lipschitz in its whole domain is rather strict, although in some circumstances such requirement is indeed met, for instance:

- *in undamped systems, the function f depends only on the system's state, therefore, the Lipschitz continuity condition is automatically verified and the differential inequation (398) becomes a differential equation, hence the control effort magnitude peters out as fast as e^{-kt}, as long as $k > 0$.*

- *in* linearly damped *systems, the transition function is of the type $f(x,v) = -\mu v + a$ term that depends on the state 'x' only. In this case, it is trivial to verify that the function f is globally L_f-Lipschitz with $L_f = \mu$.*

  In a number of applications, the design of the leader-follower systems pair is up to the user and does not have to obey any particular physical constraints (this is the case, for example, in secure transmission of information by signal masking). Therefore, linear damping may be safely assumed (which includes the case of an undamped oscillators wherever $\mu = 0$).

On the basis of the above considerations, in the context of systems on manifolds it seems reasonable to prescribe that the momentum-transition function $f : T\mathbb{M} \to T\mathbb{M}$ satisfies a mixed continuity condition in the first argument and Lipschitz continuity in the second argument. To express such requirement more precisely, let us define the following auxiliary function

$$\varphi_x(y,v) := \mathbb{P}^{y \to x}(f(y, \mathbb{P}^{x \to y}(v))), \tag{399}$$

where $x, y \in \mathbb{M}$, $v \in T_x\mathbb{M}$ and $\varphi_x(y,v) \in T_x\mathbb{M}$. Note that the following identities hold true

$$\varphi_x(x,v) = \mathbb{P}^{x \to x}(f(x, \mathbb{P}^{x \to x}(v))) = f(x,v), \tag{400}$$

$$f(x,v) - \mathbb{P}^{y \to x}(f(y,w)) = \varphi_x(x,v) - \varphi_x(x, \mathbb{P}^{y \to x}(w)), \tag{401}$$

for every $w \in T_y\mathbb{M}$. We shall say that the function (399) is uniformly continuous in the argument y and Lipschitz continuous in the argument v if for every $x \in \mathbb{M}$, for every $r, v \in T_x\mathbb{M}$ and for every $\varepsilon_\varphi > 0$ there exist a $\delta_\varphi > 0$ and a $L > 0$ such that, for each pair $z, y \in \mathbb{B}_x^\varphi$ with $\mathscr{D}(z,y) < \delta_\varphi$, it holds that

$$\|\varphi_x(y,r) - \varphi_x(z,v)\|_x < \varepsilon_\varphi + L\|r - v\|_x, \tag{402}$$

where $\mathbb{B}_x^\varphi \subset \mathbb{M}$ denotes a neighborhood of the point x. Such condition, ultimately, ensures that

$$\|f(x,v) - \mathbb{P}^{y \to x}(f(y,w))\|_x < \varepsilon_\varphi + L\|v - \mathbb{P}^{y \to x}(w)\|_x, \tag{403}$$

for every $y \in \mathbb{B}_x^\varphi$, $v \in T_x\mathbb{M}$ and $w \in T_y\mathbb{M}$, at every point $x \in \mathbb{M}$.

The above-defined mixed continuity condition is discussed in the following example concerning a specific instance of momentum-transition function.

Example 28. *Let us consider the transition function $f : T\mathbb{S}^{n-1} \to T\mathbb{S}^{n-1}$ defined by*

$$f(x,v) := -Sx + (x^\top S x)x - \mu v, \tag{404}$$

with S being a symmetric, positive-definite matrix and μ denoting a real-valued parameter. It is straightforward to verify that $(x, f(x,v))$ represents a tangent vector field in $\Gamma(\mathbb{S}^{n-1})$, in fact $x^\top f(x,v) = -x^\top S x + (x^\top S x)(x^\top x) - \mu(x^\top v) = 0$. The above function is continuous in the variable x, while it is obviously Lipschitz-continuous in the variable v.

In the following, the effects of the control term in the dynamical system (376) on the control effort magnitude in the case of a full dynamics will be analyzed in details. In order to state the main findings, it is convenient to present and comment some preliminary results.

A first instrumental result concerns the computation of the covariant derivative of a vector field obtained from the logarithmic map applied to a pair of (nearby) curves. Let $x = x(t)$ and $y = y(t)$ denote two given smooth curves on the Riemannian manifold $\mathbb{M}$ endowed with the Levi-Civita connection ∇. Consider the parametrized smooth vector field on $T\mathbb{M}$ given by $t \mapsto (x(t), \lambda(t))$, with $\lambda(t) := \log_{x(t)} y(t)$. The covariant derivative of the vector field λ along the velocity field associated to the curve x may be proven to take the expression

$$\nabla_{\dot{x}}\lambda = (d_v^{-1}\exp)_{x,\lambda}(\dot{y} - (d_x\exp)_{x,\lambda}(\dot{x})), \tag{405}$$

where $\mathrm{d}_x \exp$ and $\mathrm{d}_v \exp$ denote the partial pushforward maps of the exponential map computed with respect to its positional argument x and its tangential argument v, respectively. In order to explain such result, let a function $q : T\mathbb{M} \to \mathbb{M}$ be defined as $q(x, v) := \exp_x(v)$, with $x \in \mathbb{M}$ and $v \in T_x\mathbb{M}$. Let us denote the pushforward map of the function q with respect to the first argument by $\mathrm{d}_x q$ and the pushforward map of q with respect to the second argument by $\mathrm{d}_v q$. Since the function q is a local diffeomorphism in $T\mathbb{M}$ around $v = 0$, in a neighborhood of zero the linear map $\mathrm{d}_v q$ is full-rank, which ensure that there exists its inverse $\mathrm{d}_v^{-1} q$. In addition, notice that the identity $y(t) = \exp_{x(t)}(\log_{x(t)} y(t)) = \exp_{x(t)}(\lambda(t)) = q(x(t), \lambda(t))$ holds. Differentiating both sides with respect to the parameter t gives $\dot{y} = (\mathrm{d}_x q)_{x,\lambda}(\dot{x}) + (\mathrm{d}_v q)_{x,\lambda}(\nabla_{\dot{x}}\lambda)$, therefore, by inversion, the result relationship (405) is immediately obtained.

In order to get the reader acquainted with the notion of partial pushforwards of the exponential map, detailed computations in the case of a unit hyper-sphere are carried out in the following example.

Example 29. *Let us recall that the exponential map for the unit hypersphere corresponding to a Euclidean metric is given by*

$$\exp_x(v) = x \cos(\|v\|) + v \operatorname{sinc}(\|v\|), \tag{406}$$

where $\|\cdot\|$ denotes the standard 2-norm and 'sinc' denotes the cardinal sine function. Its partial pushforward maps may be represented as linear operators mapping $\mathbb{R}^n$ to itself, that is to say, by square matrices.

The computation of the sought tangent maps may be worked out by defining the curve

$$y(t) = \exp_{x(t)}(v(t)) = x(t) \cos(\|v(t)\|) + v(t) \operatorname{sinc}(\|v(t)\|), \tag{407}$$

where $x(t) : t \geqslant 0 \mapsto x \in \mathbb{M}$, $y(t) : t \geqslant 0 \mapsto y \in \mathbb{M}$, and $v(t) : t \geqslant 0 \mapsto v \in T_{x(t)}\mathbb{M}$. Taking the derivative of both sides of the expression (407) with respect to the parameter t yields

$$
\begin{aligned}
\dot{y} &= \dot{x}\cos(\|v\|) - x\sin(\|v\|)\frac{\mathrm{d}\|v\|}{\mathrm{d}t} + \dot{v}\operatorname{sinc}(\|v\|) + v \operatorname{sinc}'(\|v\|)\frac{\mathrm{d}\|v\|}{\mathrm{d}t} \\
&= \dot{x}\cos(\|v\|) + \dot{v}\operatorname{sinc}(\|v\|) + (\operatorname{sinc}'(\|v\|)v - x\sin(\|v\|))\frac{v^\top \dot{v}}{\|v\|},
\end{aligned}
\tag{408}
$$

where $\dot{v}$ denotes the naïve derivative of the vector field $v(t)$ and sinc' denotes the derivative of the cardinal sine function.

Recall, from Equation (115), that $\dot{v} = \nabla_{\dot{x}}v - x\dot{x}^\top v$, therefore the above expression may be recast as

$$\dot{y} = \dot{x}\cos(\|v\|) + (\nabla_{\dot{x}}v - x\dot{x}^\top v)\operatorname{sinc}(\|v\|) + (\operatorname{sinc}'(\|v\|)v - x\sin(\|v\|))\frac{v^\top \nabla_{\dot{x}}v}{\|v\|}, \tag{409}$$

because $x^\top v = 0$. A comparison of the above expression with $\dot{y} = (\mathrm{d}_x \exp)_{x,v}(\dot{x}) + (\mathrm{d}_v \exp)_{x,v}(\nabla_{\dot{x}}v)$ leads to a matrix-type representation of the sought partial pushforward maps as

$$[\![(\mathrm{d}_x \exp)_{x,v}]\!] = \cos(\|v\|)I_n - \operatorname{sinc}(\|v\|)xv^\top, \tag{410}$$

$$[\![(\mathrm{d}_v \exp)_{x,v}]\!] = \begin{cases} \operatorname{sinc}(\|v\|)(I_n - xv^\top) + \frac{\cos(\|v\|) - \operatorname{sinc}(\|v\|)}{\|v\|^2}vv^\top, & \text{for } v \neq 0, \\ I_n, & \text{for } v = 0, \end{cases} \tag{411}$$

where the symbol $[\![\cdot]\!]$ denotes a matrix representative of a linear operator.

In correspondence of the null vector $v = 0$, the pushforward map $\mathrm{d}_v \exp$ equals the identity and in a neighborhood of zero such pushforward map is invertible. For $v \neq 0$, since the matrix $[\![(\mathrm{d}_v \exp)_{x,v}]\!]$ is written as the sum of two terms, one may utilize Woodbury's matrix inversion lemma [53] to compute such inverse pushforward map. Notice that an application of the matrix

inversion lemma gives $(I_n - xv^\top)^{-1} = I_n + xv^\top$. *Applying such lemma a second time yields the sought result*

$$[\![(d_v^{-1} \exp)_{x,v}]\!] = \begin{cases} \frac{I_n + xv^\top}{\operatorname{sinc}(\|v\|)} \left[I_n - \frac{\cos(\|v\|) - \operatorname{sinc}(\|v\|)}{\|v\|^2 \cos(\|v\|)} vv^\top \right], & \text{for } v \neq 0 \\ I_n, & \text{for } v = 0. \end{cases} \tag{412}$$

Notice that a singularity occurs in correspondence of $\|v\| = \frac{\pi}{2}$, *where it holds that* $[\![(d_v \exp)_{x,v}]\!]|_{\|v\|=\pi/2} = I_n - \left(x + \frac{8}{\pi^3} v \right) v^\top$, *whose rank is 2. A second singularity occurs when* $\|v\| = \pi$, *where* $[\![(d_v \exp)_{x,v}]\!]|_{\|v\|=\pi} = -\frac{1}{\pi^2} vv^\top$, *whose rank is 1. Because of rank deficiency, in correspondence of these balls in* $T_x\mathbb{M}$ *such partial pushforward map is singular.*

The inverse map $(d_v^{-1} \exp)_{x,\log_x y} : T_y\mathbb{M} \to T_x\mathbb{M}$ plays a central role in the computation of the covariant derivative of the function $\log_x y$. Likewise, the map $(d_x \exp)_{x,\log_x y} : T_x\mathbb{M} \to T_y\mathbb{M}$ will play a central role in the analysis of the control effort dynamics. Notice that whenever $x = y$, the partial pushforward map $(d_v \exp)_{x,\log_x y}$ coincides to the identity $\operatorname{id}_{T_y\mathbb{M}}$, while the map $(d_x \exp)_{x,0}$ coincides to the identity $\operatorname{id}_{T_x\mathbb{M}}$. We are going to assume some sort of continuity of the dicussed partial pushforward maps with respect to their positional and vectorial arguments. In order to make such assumptions precise, let us define two further auxiliary maps that play a role in the computation of the covariant derivative of the logarithmic-map-based vector field, namely:

$$\mu_x(y, r) := ((d_v^{-1} \exp)_{x,\log_x y} \circ \mathbb{P}^{x \to y})(r), \tag{413}$$

$$\beta_x(y, v) := (\mathbb{P}^{y \to x} \circ (d_x \exp)_{x,\log_x y})(v), \tag{414}$$

where $x, y \in \mathbb{M}$ and $r, v \in T_x\mathbb{M}$. We chose the above notation to underline that the functions μ and β exhibit an asymmetric dependence from the variables x and y; in addition, one such notation makes it easy to remember that, for every given pair $(x, y) \in \mathbb{M}^2$, the operators $\mu_x(y, \star)$ and $\beta_x(y, \star)$ are linear endomorphisms of $T_x\mathbb{M}$. The following identities are immediate to verify:

$$\mu_x(x, r) = ((d_v^{-1} \exp)_{x,0} \circ \mathbb{P}^{x \to x})(r) = r, \tag{415}$$

$$\beta_x(x, v) = (\mathbb{P}^{x \to x} \circ (d_x \exp)_{x,0})(v) = v, \tag{416}$$

$$-\nabla_{\dot{x}} \log_x y = \mu_x(y, \beta_x(y, v) - \mathbb{P}^{y \to x}(w)). \tag{417}$$

In the following we shall assume the function μ_x defined in (413) to be continuous with respect to the variable y, namely that for every $x \in \mathbb{M}$, for every $r \in T_x\mathbb{M}$ and for every $\varepsilon_\mu > 0$, there exists a constant $\delta_\mu > 0$ such that, for any $y \in \mathbb{B}_x^\mu$ that satisfies $\mathscr{D}(x, y) < \delta_\mu$, it holds that $\|\mu_x(y, r) - r\|_x < \varepsilon_\mu$, where $\mathbb{B}_x^\mu$ denotes a neighborhood of the point $x \in \mathbb{M}$. Such assumption tells that the linear map $\mu_x(y, \star)$, thought of as a function of y, behaves like an approximate identity when x gets close to y. Note that, because of the linearity in the argument r, it is not necessary to make any restriction on its value. The continuity assumption carries over an useful consequence. By virtue of the triangular inequality, it holds that $\|\mu_x(y, r) - r\|_x > \|\mu_x(y, r)\|_x - \|r\|_x$, therefore the assumption $\|\mu_x(y, r) - r\|_x < \varepsilon_\mu$ implies that

$$\|\mu_x(y, r)\|_x < \varepsilon_\mu + \|r\|_x \text{ in } \mathbb{B}_x^\mu. \tag{418}$$

Likewise, the map $\beta_x(y, \star)$ will be assumed to be continuous with respect to the variable y, in a neighborhood $\mathbb{B}_x^\beta$ of a point x for every value $x \in \mathbb{M}$. That is to say, for every point $x \in \mathbb{M}$, every $r \in T_x\mathbb{M}$, and every scalar $\varepsilon_\beta > 0$ there exists a scalar $\delta_\beta > 0$ such that, for any $y \in \mathbb{B}_x^\beta$ with $\mathscr{D}(x, y) < \delta_\beta$, it holds that $\|\beta_x(y, v) - v\|_x < \varepsilon_\beta$.

By virtue of the triangular inequality of the norm, taking any tangent vector $w \in T_y\mathbb{M}$, we may write that $\|\beta_x(y,v) - \mathbb{P}^{y\to x}(w) - (v - \mathbb{P}^{y\to x}(w))\|_x > \|\beta_x(y,v) - \mathbb{P}^{y\to x}(w)\|_x - \|v - \mathbb{P}^{y\to x}(w)\|_x$. As a consequence, the assumption $\|\beta_x(y,v) - v\|_x < \varepsilon_\beta$ implies that

$$\|\beta_x(y,v) - \mathbb{P}^{y\to x}(w)\|_x < \varepsilon_\beta + \|v - \mathbb{P}^{y\to x}(w)\|_x \text{ in } \mathbb{B}_x^\beta. \tag{419}$$

The implications (418) and (419), together with the identity (417), tell that irrespective of how the positive values ε_μ and ε_β are chosen small, there certainly exists a point $y \in \mathbb{B}_x^\mu \cap \mathbb{B}_x^\beta \setminus \{x\}$ such that

$$\|(\mathrm{d}_v^{-1}\exp)_{x,\log_x y}[(\mathrm{d}_x\exp)_{x,\log_x y}(v) - w]\|_x < \varepsilon_{\mu|\beta} + \|v - \mathbb{P}^{y\to x}(w)\|_x, \tag{420}$$

where $\varepsilon_{\mu|\beta} := \varepsilon_\mu + \varepsilon_\beta$. The inequality (420) appears as a sort of mixed continuity/Lipschitz-continuity condition on the composition of the partial pushforward maps associated to the exponential map.

Let us evaluate the continuity of the expressions $\mu_x(y,u)$ and $\beta_x(y,v)$ for the unit-hyperpshere, as an instructive case of study, in the following example.

Example 30. *The expression of $\mu_x(y,u)$, when $x \neq y \in \mathbb{S}^{n-1}$, may be obtained by composing the expression (412) and the expression of the parallel transport on the unit hyper-sphere, that we recall to take the expression*

$$\mathbb{P}^{x\to y}(u) = \left(I_n - \frac{(x+y)y^\top}{1 + x^\top y}\right)u. \tag{421}$$

We shall avoid writing the expression of the auxiliary function μ explicitly because it is quite cumbersome and not particularly illuminating. It suffices to note that, as we have seen in the Example 29, the map (412) is continuous as long as $\|v\| < \frac{\pi}{2}$, while the map (421) is continuous as along as $\mathscr{D}(x,y) < \pi$. Therefore, the function $\mu_x(y,\star)$ is continuous in $\mathbb{B}_x^\mu := \{y \in \mathbb{S}^{n-1}|\mathscr{D}(x,y) < \frac{\pi}{2}\}$. Likewise, the expression of $\beta_x(y,v)$ may be obtained by composing the expression (410) and the expression (421), to give:

$$\beta_x(y,v) = \left(I_n - \frac{(y+x)x^\top}{1 + \cos\mathscr{D}(x,y)}\right)(I_n\cos\mathscr{D}(x,y) - xy^\top)v. \tag{422}$$

Such function is continuous in $\mathbb{B}_x^\beta := \{y \in \mathbb{S}^{n-1}|\mathscr{D}(x,y) < \pi\}$. Gluing the two above results, one gets $\mathbb{B}_x^\mu \cap \mathbb{B}_x^\beta \setminus \{x\} = \{y \in \mathbb{S}^{n-1}|0 < \mathscr{D}(x,y) < \frac{\pi}{2}\}$.

Observation 2. *The inverse map $\mu_x^{-1}(y,\star)$ has been studied in details in the unpublished contribution [54]. (In particular, to make the notation uniform, the map $\mu_x^{-1}(y,\star)$ corresponds to the map $\mathcal{E}_x(\log_x y) := \mathcal{I}_x^{-1}(\log_x y) \circ (\mathrm{d}_v\exp)_{x,\log_x y}$ in [54].) The contribution [54] shows that this map may be written explicitly as a Taylor series of higher-order covariant derivatives of the Riemannian curvature endomorphism.*

As a further instrumental result in the present analysis, we may express the covariant derivative of the parallel transport operator $\nabla\mathbb{P}$ on a manifold $\mathbb{M}$ endowed with a Levi-Civita connection ∇. Consider the parametrized smooth vector field $t \mapsto (y(t), w(t)) \in T\mathbb{M}$ and define the new parametrized vector field $t \mapsto (x(t), \mathbb{P}^{y(t)\to x(t)}(w(t))) \in T\mathbb{M}$. The covariant derivative of such vector field along the velocity field of the curve $x(t)$ takes the expression

$$\nabla_{\dot{x}}\mathbb{P}^{y\to x}(w) = (\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w}(\dot{x}) - (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x\mathbb{P}^{y\to x})_{y,x,v})(\dot{y}) + \mathbb{P}^{y\to x}(\nabla_{\dot{y}}w), \tag{423}$$

where $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w} : T_x\mathbb{M} \to T_x\mathbb{M}$ for every $y \in \mathbb{M}$ and $w \in T_y\mathbb{M}$, and $(\mathrm{d}_x\mathbb{P}^{y\to x})_{y,x,v} : T_y\mathbb{M} \to T_y\mathbb{M}$ for every $x \in \mathbb{M}$ and $v \in T_x\mathbb{M}$. In addition, the 'diagonal' part of the linear

map $(\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w}$ is identically zero, namely $(\mathrm{d}_x \mathbb{P}^{y \to x})_{x,x,v} \equiv 0$, for every point $x \in \mathbb{M}$ and tangent vector $v \in T_x \mathbb{M}$.

The above statements may be motivated as follows. Let $h : \mathbb{M} \times T\mathbb{M} \to T\mathbb{M}$ be defined as $h(x, y, w) := \mathbb{P}^{y \to x}(w)$ for any two points x and y in $\mathbb{M}$ and for any tangent vector $w \in T_y \mathbb{M}$. Note, in particular, that the function h is linear in the third argument. Now, set $v(t) := h(x(t), y(t), w(t))$ and note that v is a vector field in $\Gamma(\mathbb{M})$, namely $v(t) \in T_{x(t)} \mathbb{M}$. By definition of covariant derivative along a curve, we have that

$$
\begin{aligned}
\nabla_{\dot{x}} v &= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}^{x(t+s) \to x(t)}(v(t+s)) - v(t) \right\} \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}^{x(t+s) \to x(t)}(h(x(t+s), y(t+s), w(t+s))) - h(x(t), y(t), w(t)) \right\} \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}^{x(t+s) \to x(t)}(h(x(t+s), y(t+s), w(t+s))) - h(x(t), y(t+s), w(t+s)) \right. \\
&\qquad + h\left(x(t), y(t+s), w(t+s) - \mathbb{P}^{y(t) \to y(t+s)}(w(t)) \right) \\
&\qquad \left. + h\left(x(t), y(t+s), \mathbb{P}^{y(t) \to y(t+s)}(w(t)) \right) - h(x(t), y(t), w(t)) \right\}.
\end{aligned}
\tag{424}
$$

The covariant derivative $\nabla_{\dot{x}} v$ may thus be expressed as

$$
\nabla_{\dot{x}} v = (\mathrm{d}_x h)_{x,y,w}(\dot{x}) + (\mathrm{d}_y h)_{x,y,w}(\dot{y}) + h(x, y, \nabla_{\dot{y}} w),
\tag{425}
$$

where $\dot{x} \in T_x \mathbb{M}$, $\dot{y} \in T_y \mathbb{M}$ and $\nabla_{\dot{y}} w \in T_y \mathbb{M}$. Since the notation might lead to some confusion, let us specify that the symbol $\mathrm{d}_x h$ denotes the pushforward map of the function h with respect to its *first argument*, while the symbol $\mathrm{d}_y h$ denotes the pushforward map of the function h evaluated with respect to its *second argument*. Such partial pushforward maps are defined as

$$
\begin{cases}
(\mathrm{d}_x h)_{x,y,w}(\dot{x}) := \lim_{s \to 0} \frac{1}{s} \{ \mathbb{P}^{x(t+s) \to x(t)}(h(x(t+s), y(t+s), w(t+s))) \\
\qquad\qquad\qquad\qquad - h(x(t), y(t+s), w(t+s)) \}, \\
(\mathrm{d}_y h)_{x,y,w}(\dot{y}) := \lim_{s \to 0} \frac{1}{s} \left\{ h\left(x(t), y(t+s), \mathbb{P}^{y(t) \to y(t+s)}(w(t)) \right) - h(x(t), y(t), w(t)) \right\}.
\end{cases}
\tag{426}
$$

Moreover, the notation $(\mathrm{d}_x h)_{x,y,w}$ indicates that the pushforward map, that is a function of three variables (besides of its argument), is evaluated at a point where the first variable is instantiated to the value x, the second variable is instantiated to the value y and the third variable is instantiated to the value w.

Now, since $\mathbb{P}^{x \to y} \circ \mathbb{P}^{y \to x} = \mathrm{id}_{T_x \mathbb{M}}$, the identity $w = h(y, x, h(x, y, w))$ holds. Applying to both sides the covariant derivative ∇ yields the identity

$$
\nabla_{\dot{y}} w = (\mathrm{d}_x h)_{y,x,v}(\dot{y}) + (\mathrm{d}_y h)_{y,x,v}(\dot{x}) + h(y, x, (\mathrm{d}_x h)_{x,y,w}(\dot{x}) + (\mathrm{d}_y h)_{x,y,w}(\dot{y}) + h(x, y, \nabla_{\dot{y}} w)).
\tag{427}
$$

Notice that $(\mathrm{d}_x h)_{y,x,v}$ is an operator that maps $T_y \mathbb{M}$ to itself, while the operator $(\mathrm{d}_y h)_{x,y,w}$ maps $T_y \mathbb{M}$ to $T_x \mathbb{M}$. Since the velocity vector fields $\dot{x}$ and $\dot{y}$ are arbitrary, the above identity, rewritten as

$$
\nabla_{\dot{y}} w = (\mathrm{d}_x h)_{y,x,v}(\dot{y}) + (\mathrm{d}_y h)_{y,x,v}(\dot{x}) + h(y, x, (\mathrm{d}_x h)_{x,y,w}(\dot{x})) + h(y, x, (\mathrm{d}_y h)_{x,y,w}(\dot{y})) + \nabla_{\dot{y}} w,
\tag{428}
$$

yields two independent operator identities, one of which is $h(y, x, (\mathrm{d}_y h)_{x,y,w}) + (\mathrm{d}_x h)_{y,x,h(x,y,w)} = 0$, namely

$$
(\mathrm{d}_y h)_{x,y,w} = -h(x, y, (\mathrm{d}_x h)_{y,x,h(x,y,w)}).
\tag{429}
$$

By using the Equation (425) and the definition of the function h leads to the sought result.

The statement about the diagonal part of the tangent map $(\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w}$ follows from the definitions (426), in fact

$$
\begin{aligned}
(\mathrm{d}_x h)_{x,x,v}(\dot{x}) &= \lim_{s \to 0} \frac{1}{s} \{ \mathbb{P}^{x(t+s) \to x(t)}(h(x(t+s), x(t+s), v(t+s))) \\
&\qquad\qquad - h(x(t), x(t+s), v(t+s)) \}, \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}^{x(t+s) \to x(t)}(v(t+s)) - h(x(t), x(t+s), v(t+s)) \right\}, \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}^{x(t+s) \to x(t)}(v(t+s)) - \mathbb{P}^{x(t+s) \to x(t)}(v(t+s)) \right\}, \\
&= 0,
\end{aligned}
\tag{430}
$$

no matter how the point $x \in \mathbb{M}$ and the tangent vector $v \in T_x \mathbb{M}$ are chosen on the tangent bundle. But since $\dot{x}$ is arbitrary, the statement follows.

The following example clarifies the above findings.

Example 31. *Consider the parallel translation map for the unit hyper-sphere $\mathbb{S}^{n-1}$ corresponding to a Euclidean metric. By applying the definitions of partial pushforward maps in terms of time-derivatives, it is not hard to compute the maps $(\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w}$ and $(\mathrm{d}_y \mathbb{P}^{y \to x})_{x,y,w}$. Let us initiate by recalling that, in the present example,*

$$
h(x, y, w) := \left(I_n - \frac{(y+x)x^\top}{1 + x^\top y} \right) w.
\tag{431}
$$

(Notice that this instance of parallel transport is from y to x! In fact, $w \in T_y \mathbb{M}$). The naïve derivative of such instance of the function h reads

$$
\dot{h} = \left(-\frac{(\dot{y}+\dot{x})x^\top(1+x^\top y) + (y+x)\dot{x}^\top(1+x^\top y) - (y+x)x^\top(\dot{x}^\top y + x^\top \dot{y})}{(1+x^\top y)^2} \right) w
\tag{432}
$$
$$
+ h(x, y, \dot{w}).
$$

By making use of the expression of the covariant derivative of the vector fields $v = h(x, y, w)$ and w given in (115), the relation (425) may be recast as:

$$
\dot{h} + x h^\top \dot{x} = (\mathrm{d}_x h)_{x,y,w}(\dot{x}) + (\mathrm{d}_y h)_{x,y,w}(\dot{y}) + h(x, y, \dot{w} + y w^\top \dot{y}).
\tag{433}
$$

Plugging the Equation (432) into the relation (433) leads to

$$
-\frac{\left((\dot{y}+\dot{x})(x^\top w)(1+x^\top y) + (y+x)(\dot{x}^\top w)(1+x^\top y) - (y+x)(x^\top w)(\dot{x}^\top y + x^\top \dot{y}) \right)}{(1+x^\top y)^2}
$$
$$
+ x w^\top \left(I_n - \frac{x(y+x)^\top}{1+x^\top y} \right) \dot{x} = (\mathrm{d}_x h)_{x,y,w}(\dot{x}) + (\mathrm{d}_y h)_{x,y,w}(\dot{y}) + (w^\top \dot{y}) \left(I_n - \frac{(y+x)x^\top}{1+x^\top y} \right) y.
\tag{434}
$$

Upon splitting the terms in $\dot{x}$ and $\dot{y}$, after some tedious but not hard calculations, we obtain a matrix representation of the sought partial pushforward maps given by

$$
[\![(\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w}]\!] = -\frac{(x^\top w)(I_n + xy^\top) + (y+x)w^\top}{1+x^\top y} + x w^\top + \frac{(x^\top w)(y+x)y^\top}{(1+x^\top y)^2},
\tag{435}
$$

$$
[\![(\mathrm{d}_y \mathbb{P}^{y \to x})_{x,y,w}]\!] = -\frac{x^\top w}{1+x^\top y} \left(I_n - \frac{(y+x)x^\top}{1+x^\top y} \right) - y w^\top + \frac{(x^\top y)(y+x)w^\top}{1+x^\top y}.
\tag{436}
$$

It is not hard to verify that $x^\top (\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}(\dot{x}) = x^\top (\mathrm{d}_y \mathbb{P}^{y\to x})_{x,y,w}(\dot{y}) = 0$ by using the above matrix representations. In addition, some lengthy – yet straightforward – calculations allow one to verify the identity (429), that may be rewritten explicitly as

$$(\mathrm{d}_y \mathbb{P}^{y\to x})_{x,y,w}(\dot{y}) = -\mathbb{P}^{y\to x} \circ (\mathrm{d}_x \mathbb{P}^{y\to x})_{y,x,\mathbb{P}^{y\to x}(w)}(\dot{y}). \tag{437}$$

Such identity establishes that the partial pushforward map $\mathrm{d}_y \mathbb{P}$ may be written in terms of the partial pushforward map $\mathrm{d}_x \mathbb{P}$, which justifies why the latter is termed principal pushforward map *(of the parallel transport operator). It is also straightforward to verify directly that $\llbracket (\mathrm{d}_x \mathbb{P}^{y\to x})_{x,x,v} \rrbracket = 0$ for the diagonal part of the partial pushforward map $\mathrm{d}_x \mathbb{P}$.*

Some remarks on the partial pushforward maps of the parallel transport operator are in order. The map $(\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}$ is linear in the variable w (as well as in its argument), therefore, the only way it could possibly be bounded over the whole domain $\mathbb{M} \times T\mathbb{M}$ is that it is identically zero. Moreover, since the linear map $(\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}$ gets close to the null map when the points x and y come close to one another, it is natural to express the continuity of the term $(\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}(\dot{x}) - (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x \mathbb{P}^{y\to x})_{y,x,v})(\dot{y})$ by requiring that the function $(\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}(v) - (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x \mathbb{P}^{y\to x})_{y,x,\mathbb{P}^{y\to x}(w)})(\tau)$ is continuous within its domain of definition if, for every $\varepsilon_\psi > 0$, there exists a scalar $\delta_\psi > 0$ such that, for any pair $x, y \in \mathbb{M}^2$ that satisfies the inequality $\mathscr{D}(x,y) < \delta_\psi$, it results that

$$\| (\mathrm{d}_x \mathbb{P}^{y\to x})_{x,y,w}(v) - (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x \mathbb{P}^{y\to x})_{y,x,\mathbb{P}^{y\to x}(w)})(\tau) \|_x < \varepsilon_\psi, \tag{438}$$

for every choice of the tangent vectors $w, \tau \in T_y\mathbb{M}$ and $v \in T_x\mathbb{M}$.

We have by now set all the analytic tools that are indeed necessary to justify the main result of the present section. We shall in fact show that the control scheme (376) ensures that the control effort vanishes to zero under conditions of continuity.

Let us assume that the parallel-transport operator and the manifold exponential map are locally continuous together with their tangent maps in the tangent bundle $T\mathbb{M}$. Let us further assume that the velocity-transition function is continuous in its first argument and Lipschitz continuous in its second argument with Lipschitz constant L_f. If the constant k meets the conditions $k \neq L_f + \frac{c}{4}$ and $k > L_f$, then the control effort presents a convergence to zero that results to be almost exponentially fast.

In order to justify such statement, it is convenient for easier tractability to define a *dual control field* $(x_\mathrm{m}, u^\star) \in T\mathbb{M}$ as

$$u^\star = v_\mathrm{m} - \mathbb{P}^{x_\mathrm{s} \to x_\mathrm{m}}(v_\mathrm{s}) - \tfrac{c}{2} \log_{x_\mathrm{m}}(x_\mathrm{s}). \tag{439}$$

It is worth noticing that, since $u^\star = \mathbb{P}^{x_\mathrm{s} \to x_\mathrm{m}}(u)$, it holds that

$$\sigma = \tfrac{1}{2} \|u\|_{x_\mathrm{s}}^2 = \tfrac{1}{2} \|u^\star\|_{x_\mathrm{m}}^2. \tag{440}$$

Moreover, since the manifold $\mathbb{M}$ was endowed with a metric connection ∇, the control effort possesses a time-dynamics described by the covariant derivative of the dual control field through the expression

$$\dot{\sigma} = \langle u^\star, \nabla_{\dot{x}_\mathrm{m}} u^\star \rangle_{x_\mathrm{m}}. \tag{441}$$

The covariant derivative of the dual control field, in turn, consists of three terms:

$$\nabla_{\dot{x}_\mathrm{m}} u^\star = \nabla_{\dot{x}_\mathrm{m}} v_\mathrm{m} - \nabla_{\dot{x}_\mathrm{m}} \mathbb{P}^{x_\mathrm{s} \to x_\mathrm{m}}(v_\mathrm{s}) - \tfrac{c}{2} \nabla_{\dot{x}_\mathrm{m}} \log_{x_\mathrm{m}}(x_\mathrm{s}). \tag{442}$$

According to the expression (405), the covariant derivative of the logarithmic map applied to the trajectories of the leader and of the follower subsystems reads

$$\nabla_{\dot{x}_\mathrm{m}} \log_{x_\mathrm{m}}(x_\mathrm{s}) = (\mathrm{d}_v^{-1} \exp)_{x_\mathrm{m}, \log_{x_\mathrm{m}}(x_\mathrm{s})} (\dot{x}_\mathrm{s} - (\mathrm{d}_x \exp)_{x_\mathrm{m}, \log_{x_\mathrm{m}}(x_\mathrm{s})}(\dot{x}_\mathrm{m})). \tag{443}$$

In addition, according to the expression (423), the covariant derivative of the parallel-transport-related field reads

$$
\begin{aligned}
\nabla_{\dot{x}_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}}) &= (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{m}}, x_{\mathrm{s}}, v_{\mathrm{s}}}(\dot{x}_{\mathrm{m}}) + \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(\nabla_{\dot{x}_{\mathrm{s}}} v_{\mathrm{s}}) \\
&\quad - (\mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}} \circ (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{s}}, x_{\mathrm{m}}, \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}})})(\dot{x}_{\mathrm{s}}) \\
&= (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{m}}, x_{\mathrm{s}}, v_{\mathrm{s}}}(\dot{x}_{\mathrm{m}}) + \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(f(x_{\mathrm{s}}, v_{\mathrm{s}})) + k\, u^{\star} \\
&\quad - (\mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}} \circ (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{s}}, x_{\mathrm{m}}, \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}})})(\dot{x}_{\mathrm{s}}).
\end{aligned}
\tag{444}
$$

Putting all those terms together in a single expression gives, for the covariant derivative of the dual control field along the trajectory of the leader subsystem, that reads

$$
\begin{aligned}
\nabla_{\dot{x}_{\mathrm{m}}} u^{\star} &= f(x_{\mathrm{m}}, v_{\mathrm{m}}) - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(f(x_{\mathrm{s}}, v_{\mathrm{s}})) \\
&\quad - (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{m}}, x_{\mathrm{s}}, v_{\mathrm{s}}}(\dot{x}_{\mathrm{m}}) + (\mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}} \circ (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{s}}, x_{\mathrm{m}}, \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}})})(\dot{x}_{\mathrm{s}}) \\
&\quad - \tfrac{c}{2} (\mathrm{d}_v^{-1} \exp)_{x_{\mathrm{m}}, \log_{x_{\mathrm{m}}}(x_{\mathrm{s}})}(\dot{x}_{\mathrm{s}} - (\mathrm{d}_x \exp)_{x_{\mathrm{m}}, \log_{x_{\mathrm{m}}}(x_{\mathrm{s}})}(\dot{x}_{\mathrm{m}})) \\
&\quad - k\, u^{\star}.
\end{aligned}
\tag{445}
$$

It is convenient to recall from the very definition of dual control field that

$$
\begin{cases}
u^{\star} + \tfrac{c}{2} \log_{x_{\mathrm{m}}}(x_{\mathrm{s}}) = v_{\mathrm{m}} - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}}), \\
\dot{x}_{\mathrm{m}} - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(\dot{x}_{\mathrm{s}}) = v_{\mathrm{m}} - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}}) - u^{\star},
\end{cases}
\tag{446}
$$

from which, by the triangular property of the norm, it follows directly that

$$
\begin{cases}
\| v_{\mathrm{m}} - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}}) \|_{x_{\mathrm{m}}} \leqslant \| u^{\star} \|_{x_{\mathrm{m}}} + \tfrac{c}{2} \mathscr{D}(x_{\mathrm{s}}, x_{\mathrm{m}}), \\
\| \dot{x}_{\mathrm{m}} - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(\dot{x}_{\mathrm{s}}) \|_{x_{\mathrm{m}}} = \tfrac{c}{2} \mathscr{D}(x_{\mathrm{s}}, x_{\mathrm{m}}).
\end{cases}
\tag{447}
$$

Let us also recall that the distance $\mathscr{D}(x_{\mathrm{s}}(t), x_{\mathrm{m}}(t))$ takes on the expression $\mathscr{D}_0 \exp\left(-\tfrac{c}{2} t\right)$, where $\mathscr{D}_0 := \mathscr{D}(x_{\mathrm{s},0}, x_{\mathrm{m},0})$.

For the inner product in Equation (441), it holds that

$$
\begin{aligned}
\langle u^{\star}, \nabla_{\dot{x}_{\mathrm{m}}} u^{\star} \rangle_{x_{\mathrm{m}}} =\ & \\
& - k \langle u^{\star}, u^{\star} \rangle_{x_{\mathrm{m}}} \\
& + \langle u^{\star}, f(x_{\mathrm{m}}, v_{\mathrm{m}}) - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(f(x_{\mathrm{s}}, v_{\mathrm{s}})) \rangle_{x_{\mathrm{m}}} \\
& + \langle u^{\star}, (\mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}} \circ (\mathrm{d}_{z_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{s}}, x_{\mathrm{m}}, \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(v_{\mathrm{s}})})(\dot{x}_{\mathrm{s}}) - (\mathrm{d}_{x_{\mathrm{m}}} \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}})_{x_{\mathrm{m}}, x_{\mathrm{s}}, v_{\mathrm{s}}}(\dot{x}_{\mathrm{m}}) \rangle_{x_{\mathrm{m}}} \\
& + \tfrac{c}{2} \langle u^{\star}, (\mathrm{d}_v^{-1} \exp)_{x_{\mathrm{m}}, \log_{x_{\mathrm{m}}}(x_{\mathrm{s}})}((\mathrm{d}_x \exp)_{x_{\mathrm{m}}, \log_{x_{\mathrm{m}}}(x_{\mathrm{s}})}(\dot{x}_{\mathrm{m}}) - \dot{x}_{\mathrm{s}}) \rangle_{x_{\mathrm{m}}}.
\end{aligned}
\tag{448}
$$

Let us examine the four terms in the right-hand side of the above lengthy expression:

- the term $-k \langle u^{\star}, u^{\star} \rangle_{x_{\mathrm{m}}}$ clearly equals $-2k\,\sigma$ because of the identity (440);
- the term $\langle u^{\star}, f(x_{\mathrm{m}}, v_{\mathrm{m}}) - \mathbb{P}^{x_{\mathrm{s}} \to x_{\mathrm{m}}}(f(x_{\mathrm{s}}, v_{\mathrm{s}})) \rangle_{x_{\mathrm{m}}}$ may be majorized by means of the Cauchy-Schwarz inequality, by the assumption that the function f is continuous in the first argument and Lipschitz-continuous in the second argument, and by the properties (447). Namely, $\forall \varepsilon_\varphi > 0, \exists \delta_\varphi > 0, L_f > 0$ such that $\forall x_{\mathrm{m}}, x_{\mathrm{s}} \in \mathbb{M}$ with $\mathscr{D}(x_{\mathrm{m}}, x_{\mathrm{s}}) < \delta_\varphi$, it holds that

$$
\begin{aligned}
\langle u^\star, f(x_{\mathrm m}, v_{\mathrm m}) &- \mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(f(x_{\mathrm s}, v_{\mathrm s}))\rangle_{x_{\mathrm m}} \\
&\leqslant \|u^\star\|_{x_{\mathrm m}}\|f(x_{\mathrm m}, v_{\mathrm m}) - \mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(f(x_{\mathrm s}, v_{\mathrm s}))\| \\
&< \|u^\star\|_{x_{\mathrm m}}\Big(\varepsilon_\varphi + L_f\|v_{\mathrm m} - \mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(v_{\mathrm s})\|_{x_{\mathrm m}}\Big) \\
&\leqslant \|u^\star\|_{x_{\mathrm m}}\Big(\varepsilon_\varphi + L_f(\|u^\star\|_{x_{\mathrm m}} + \tfrac{c}{2}\mathscr{D}(x_{\mathrm s}, x_{\mathrm m}))\Big) \\
&= L_f\|u^\star\|_{x_{\mathrm m}}^2 + \|u^\star\|_{x_{\mathrm m}}\Big(\varepsilon_\varphi + \tfrac{1}{2}L_f c \mathscr{D}(x_{\mathrm s}, x_{\mathrm m})\Big) \\
&= 2L_f\sigma + \sqrt{2\sigma}\Big(\varepsilon_\varphi + \tfrac{1}{2}L_f c \mathscr{D}(x_{\mathrm s}, x_{\mathrm m})\Big);
\end{aligned}
\tag{449}
$$

- the further scalar component $\langle u^\star, (\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}} \circ (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm s},x_{\mathrm m},\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(v_{\mathrm s})})(\dot{x}_{\mathrm s}) - (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm m},x_{\mathrm s},v_{\mathrm s}}(\dot{x}_{\mathrm m})\rangle_{x_{\mathrm m}}$ may be majorized by invoking the Cauchy-Schwarz inequality and the continuity of the parallel transport and of its tangent maps. Namely, $\forall \varepsilon_\psi > 0, \exists \delta_\psi > 0$ such that $\forall x_{\mathrm m}, x_{\mathrm s} \in \mathbb{M}$ with $\mathscr{D}(x_{\mathrm m}, x_{\mathrm s}) < \delta_\psi$, it holds that

$$
\begin{aligned}
\langle u^\star, (\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}} &\circ (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm s},x_{\mathrm m},\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(v_{\mathrm s})})(\dot{x}_{\mathrm s}) - (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm m},x_{\mathrm s},v_{\mathrm s}}(\dot{x}_{\mathrm m})\rangle_{x_{\mathrm m}} \\
&\leqslant \|u^\star\|_{x_{\mathrm m}}\|(\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}} \circ (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm s},x_{\mathrm m},\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(v_{\mathrm s})})(\dot{x}_{\mathrm s}) - (\mathrm{d}_{x_{\mathrm m}}\mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}})_{x_{\mathrm m},x_{\mathrm s},v_{\mathrm s}}(\dot{x}_{\mathrm m})\|_{x_{\mathrm m}} \\
&< \varepsilon_\psi\|u^\star\|_{x_{\mathrm m}} \\
&= \sqrt{2\sigma}\varepsilon_\psi;
\end{aligned}
\tag{450}
$$

- to end with, the term $\frac{c}{2}\langle u^\star, (\mathrm{d}_v^{-1}\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}((\mathrm{d}_x\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}(\dot{x}_{\mathrm m}) - \dot{x}_{\mathrm s})\rangle_{x_{\mathrm m}}$ may be majorized by the help of the Cauchy-Schwarz inequality, by the assumptions on the continuity of the pushforward maps of the exponential map and by the inequalities (447), namely, $\forall \varepsilon_\beta > 0, \exists \delta_\beta > 0$ such that $\forall x_{\mathrm s}, x_{\mathrm m} \in \mathbb{M}$ for which $\mathscr{D}(x_{\mathrm s}, x_{\mathrm m}) < \delta_\beta$ then

$$
\begin{aligned}
\tfrac{c}{2}\langle u^\star, (\mathrm{d}_v^{-1}\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}&((\mathrm{d}_x\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}(\dot{x}_{\mathrm m}) - \dot{x}_{\mathrm s})\rangle_{x_{\mathrm m}} \\
&\leqslant \tfrac{c}{2}\|u^\star\|_{x_{\mathrm m}}\|(\mathrm{d}_v^{-1}\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}((\mathrm{d}_x\exp)_{x_{\mathrm m},\log_{x_{\mathrm m}}(x_{\mathrm s})}(\dot{x}_{\mathrm m}) - \dot{x}_{\mathrm s})\|_{x_{\mathrm m}} \\
&< \tfrac{c}{2}\|u^\star\|_{x_{\mathrm m}}\big(\varepsilon_\beta + \|\dot{x}_{\mathrm m} - \mathbb{P}^{x_{\mathrm s}\to x_{\mathrm m}}(\dot{x}_{\mathrm s})\|_{x_{\mathrm m}}\big) \\
&= \tfrac{c}{2}\|u^\star\|_{x_{\mathrm m}}\big(\varepsilon_\beta + \tfrac{c}{2}\mathscr{D}(x_{\mathrm s}, x_{\mathrm m})\big) \\
&= \tfrac{c}{2}\sqrt{2\sigma}\big(\varepsilon_\beta + \tfrac{c}{2}\mathscr{D}(x_{\mathrm s}, x_{\mathrm m})\big).
\end{aligned}
\tag{451}
$$

Majorizing the quantity in (448) by means of the inequalities (449)–(451), and choosing $\varepsilon_\varphi = \frac{\varepsilon}{3}, \varepsilon_\psi = \frac{\varepsilon}{3}$ and $\varepsilon_\beta = \frac{2\varepsilon}{3c}$, for t sufficiently large to fulfill $\mathscr{D}(x_{\mathrm s}, x_{\mathrm m}) < \min\{\delta_\varphi, \delta_\beta, \delta_\psi\}$, it holds that

$$
\dot{\sigma}(t) < -k^\star\sigma(t) + \sqrt{2\sigma(t)}\rho(t),
\tag{452}
$$

with $k^\star$ and $\rho(t)$ defined as in (389). By a Grönwall-like majorization argument, we conclude that the control effort vanishes to zero at least exponentially.

6.4. Feedback Error Velocity, Principal Pushforward Map and Relation to Curvature

The cornerstone of control theory is *error feedback control*, based on the definition of a suitable error figure between the actual state of a system to control and a desired state, to be fed to a control algorithm. The control algorithm, in turn, will produce the appropriate commands to the system aimed at reducing the value of the error, hence forming a closed loop. Denoting by e such control error, it pays to study the evolution law of the error, often determined by a differential equation in the terms $\dot{e}$ and $\ddot{e}$, in order to make sure that effectively the error converges to zero (asymptotically). Such differential equation describes an *error system*, as it was exemplified in Section 6.1. The above-mentioned time-derivatives of the error also intervene in designing the control algorithm itself, which in industrial applications is often chosen to include proportional, integral and derivative (PID) terms. It is, in fact, known that each of these term may contribute in shaping the error system and in

compensating each others' drawbacks in order to attain a desired dynamics for the error field e.

When considering two dynamical systems whose common state space is a curved Riemannian manifold $\mathbb{M}$, whose states are described by the manifold-valued variables $x, y \in \mathbb{M}$ (where let us say that the reference system has state-variable x), there is not a straightforward definition of error. According to the prevailing literature (see, e.g., [55,56]), and on the present author's personal experience, the error term $e \in T_x\mathbb{M}$ should be defined as a tangent field along the trajectory of one of the systems as, for example, $e := \log_x y$. A problem with this definition often comes when formulating the derivative and integral control terms. What is most relevant to the present discussion is the computation of a derivative term which, for consistency, cannot be the naïve derivative $\dot{e}$ of the error but needs to be, for example, the covariant derivative $\varepsilon := \nabla_t^x e = \nabla_t^x \log_x y$. Such calculation was already detailed in Section 6.3 and we may recall from Equation (405) that

$$\varepsilon = \nabla_{\dot{x}} \log_x y = (\mathrm{d}_v^{-1} \exp)_{x, \log_x y}(\dot{y}) - (\mathrm{d}_v^{-1} \exp)_{x, \log_x y}((\mathrm{d}_x \exp)_{x, \log_x y}(\dot{x})). \quad (453)$$

In practice, such expression, which depends on the partial pushforwards of the exponential map, is deemed to be overly convoluted to employ and hence is replaced by a simpler expression that retains the meaning of *error velocity* and the property of being a tangent vector field yet being much simpler to utilize, namely

$$\varepsilon := \mathbb{P}^{y \to x}(\dot{y}) - \dot{x}. \quad (454)$$

It is interesting to recall that such an error term appears in the Cucker-Smale model on manifold recalled in the introductory discussion of the present Section.

Now, the dynamics of the error depends on the (covariant) time-derivative of the velocity of the error, namely on $\nabla_t^x \varepsilon = \nabla_t^x \mathbb{P}^{y \to x}(\dot{y}) - \nabla_t^x \dot{x}$. The second term coincides with the state-acceleration of one of the systems, while the term $\nabla_t^x \mathbb{P}^{y \to x}(\dot{y})$ needs to be studied aside. The present section is devoted to the analysis of the quantity $\nabla_{\dot{x}} \mathbb{P}^{y \to x}(\dot{y})$. The analysis carried out in the present section elucidates some interesting algebraic properties of the principal pushforward map associated to the parallel transport operator and the connection between such principal pushforward map and the curvature endomorphism of the state manifold.

The scientific content recalled in the present section is largely based on the earlier contribution [18].

6.4.1. Algebraic Properties of the Principal Pushforward Map of Parallel Transport

We have already shown in Section 6.3 and, in particular, by Equation (423) that, given two smooth curves $t \mapsto x(t)$ and $t \mapsto y(t)$ in $\mathbb{M}$ and a smooth vector field $w_{y(t)} \in T_{y(t)}\mathbb{M}$, the covariant derivative of the vector field $x \mapsto \mathbb{P}^{y \to x}(w_y)$ along the velocity field of the curve $x(t)$ reads

$$\nabla_{\dot{x}} \mathbb{P}^{y \to x}(w_y) = (\mathrm{d}_x \mathbb{P}^{y \to x})_{x, y, w_y}(\dot{x}) - \mathbb{P}^{y \to x}((\mathrm{d}_x \mathbb{P}^{y \to x})_{y, x, \mathbb{P}^{y \to x}(w_y)})(\dot{y})) + \mathbb{P}^{y \to x}(\nabla_{\dot{y}} w_y), \quad (455)$$

where $(\mathrm{d}_x \mathbb{P})_{x, \star, \star} \in \mathrm{End}(T_x\mathbb{M})$ for any $x \in \mathbb{M}$. Such differential map is termed *principal pushforward map* since the covariant derivative $\nabla \mathbb{P}$ may be expressed entirely on the basis of such map without any need to refer to the further partial pushforward map $\mathrm{d}_v \mathbb{P}$.

One might notice that the covariant derivative $\nabla_{\dot{x}} \mathbb{P}^{y \to x}(w_y)$ of the transported field equals the sum of an 'obvious' term, given by the transported covariant derivative $\mathbb{P}^{y \to x}(\nabla_{\dot{y}} w_y)$, and of further terms purely related to the dynamics of parallel transport and, hence, on the curvature of the state manifold $\mathbb{M}$. (Such terms appear to be somehow reminiscent of the kinematic corrections due to a change in reference systems in physics.)

The principal pushforward map enjoys some interesting algebraic properties. Since these are not directly instrumental to the present analysis, let us recall them under the form of an example.

Example 32. *Denoting by* $(\mathrm{d}_x\mathbb{P}^{y\to x})^\dagger_{x,y,w_y} : T_x\mathbb{M} \to T_x\mathbb{M}$ *the adjoint (or dual) of the operator* $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}$ *with respect to the Riemannian metric of* $\mathbb{M}$, *it holds that*

$$(\mathrm{d}_x\mathbb{P}^{y\to x})^\dagger_{x,y,\mathbb{P}^{x\to y}(v_x)}(v_x) = 0. \tag{456}$$

for every $x,y \in \mathbb{M}$ *and* $v_x \in T_x\mathbb{M}$. *In order to explain such statement, let us recall that, since the parallel transport is an isometry, the vector fields* $t \mapsto (x(t), v(t))$ *and* $t \mapsto (y(t), w(t))$ *enjoy the property*

$$\|v_x(t)\|^2_{x(t)} = \|w_y(t)\|^2_{y(t)}. \tag{457}$$

Taking the derivative of both sides with respect to the parameter t yields the equation

$$\langle v_x, \nabla_{\dot{x}}v_x\rangle_x = \langle w_y, \nabla_{\dot{y}}w_y\rangle_y. \tag{458}$$

By replacing the expression (455) for the covariant derivative $\nabla_{\dot{x}}v_x$ *in (458), we get:*

$$\langle v_x, (\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_x}(\dot{x})\rangle_x - \langle v_x, (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x\mathbb{P}^{y\to x})_{y,x,v_x})(\dot{y})\rangle_x$$
$$+ \langle \mathbb{P}^{y\to x}(w_y), \mathbb{P}^{y\to x}(\nabla_{\dot{y}}w_y)\rangle_x = \langle w_y, \nabla_{\dot{y}}w_y\rangle_y.$$

By way of the isometry property of parallel transport, it is immediate to see that the last term on the left-hand side equals the term on the right-hand side, therefore the above property is equivalent to

$$\langle v_x, (\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}(\dot{x})\rangle_x = \langle v_x, (\mathbb{P}^{y\to x} \circ (\mathrm{d}_x\mathbb{P}^{y\to x})_{y,x,v_x})(\dot{y})\rangle_x. \tag{459}$$

It pays to recall that, given a linear operator $L : \mathbb{V} \to \mathbb{V}$ *on a vector space* $\mathbb{V}$ *endowed with a positive-definite inner product* $\langle\cdot,\cdot\rangle : \mathbb{V}^2 \to \mathbb{R}$, *its adjoint* $L^\dagger$ *with respect to the metric satisfies the relation* $\langle L(v), u\rangle = \langle v, L^\dagger(u)\rangle$ *for every* $u, v \in \mathbb{V}$. *The left-hand side of the Equation (459) may thus be rewritten as* $\langle(\mathrm{d}_x\mathbb{P}^{y\to x})^\dagger_{x,y,\mathbb{P}^{x\to y}(v_x)}(v_x), \dot{x}\rangle_x$. *Since* $x, y, \dot{x}$ *and* $\dot{y}$ *are arbitrary, the Equation (459) may hold only if the identity (456) holds.*

Let us further recall the following two properties of linear operators on vector spaces:

- ***Equivalence of matrix representations***: *Given a linear operator* $L : \mathbb{V} \to \mathbb{V}$ *on a vector space* $\mathbb{V} \subset \mathbb{R}^m$, *two of its matrix representations* $[\![L]\!]_1$ *and* $[\![L]\!]_2$ *are equivalent on the vector space* $\mathbb{V}$ *if* $[\![L]\!]_1 x = [\![L]\!]_2 x$ *for any* $x \in \mathbb{V}$, *although the matrices* $[\![L]\!]_1$ *and* $[\![L]\!]_2$ *may appear different.*
- ***Representation of the adjoint***: *If* $[\![L]\!]$ *is a matrix representation of a linear operator* $L : \mathbb{V} \to \mathbb{V}$ *on a vector space* $\mathbb{V} \subset \mathbb{R}^m$ *with respect to an orthonormal basis of* $\mathbb{V}$, *then the matrix representation of its adjoint, namely* $[\![L^\dagger]\!]$ *coincides with the transpose* $[\![L]\!]^T$ *(this is a reason why the adjoint is sometimes also referred to as* transpose operator*).*

It is immediate to recognize that the matrix-representation (435) of the operator $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}$ *is* not *unique. For example, if we add to the matrix* $[\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}]\!]$ *the term* $x^\top w_y(1 + x^\top y)^{-1}xx^\top$, *we get an equivalent representation of the map* $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}$ *given by:*

$$\begin{aligned}[\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}]\!]_1 :=& -\frac{(x^\top w_y)(I_n + xy^\top - xx^\top) + (y+x)w_y^\top}{1 + x^\top y}\\ &+ xw_y^\top + \frac{(x^\top w_y)(y+x)y^\top}{(1 + x^\top y)^2}.\end{aligned} \tag{460}$$

A further equivalent representation is

$$[\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}]\!]_2 := [\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}]\!]_1 - \frac{(w_y^\top x)^2}{(1+x^\top y)^2}\frac{[\mathbb{P}^{y\to x}(w)]x^\top}{\|w_y\|^2}, \tag{461}$$

that serves to exemplify the property (456). In fact, in the present context, it holds that $[\![(\mathrm{d}_x\mathbb{P}^{y\to x})^\dagger_{x,y,w_y}]\!] = [\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}]\!]^\top$ *and it is not hard to verify that*

$[\![(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,\mathbb{P}^{y\to x}(v_x)}]\!]_2^\top v_x = 0$, *no matter how one chooses the points $x \in \mathbb{S}^{n-1}$, $y \in \mathbb{S}^{n-1} \setminus \{-x\}$ and the tangent vector v in $T_x\mathbb{S}^{n-1}$.*

As a further property of the principal pushforward map, more related to system dynamics, let us consider two trajectories $t \mapsto x(t)$ and $t \to y(t)$ that intersect at a point $p \in \mathbb{M}$. Since $v_x = \mathbb{P}^{y\to x}(w_y)$, when $x = y = p$ the vector fields v_x and w_y align to one another at p, namely $v_p \equiv w_p \in T_p\mathbb{M}$. On the other hand, the two curves are arbitrary and we assume that $\dot{x}_p \neq \dot{y}_p$. An interesting finding is that at the point p, $\nabla_{\dot{x}}v_x = \nabla_{\dot{y}}w_y$. In fact, at any points $x, y \in \mathbb{M}$ and for every $w_y \in T_y\mathbb{M}$, it holds that

$$(\nabla_{\dot{x}}v_x)_x = (\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}(\dot{x}_x) - \mathbb{P}^{y\to x}\Big((\mathrm{d}_x\mathbb{P}^{y\to x})_{y,x,\mathbb{P}^{y\to x}(w_y)}(\dot{y}_y) - (\nabla_{\dot{y}}w_y)_y\Big). \quad (462)$$

Setting $x = y = p$ gives

$$(\nabla_{\dot{x}}v_x)_p = (\mathrm{d}_x\mathbb{P}^{y\to x})_{p,p,w_p}(\dot{x}_p) - (\mathrm{d}_x\mathbb{P}^{y\to x})_{p,p,w_p}(\dot{y}_p) + (\nabla_{\dot{y}}w_y)_p, \quad (463)$$

namely

$$(\nabla_{\dot{x}}v_x)_p - (\nabla_{\dot{y}}w_y)_p = (\mathrm{d}_x\mathbb{P}^{y\to x})_{p,p,w_p}(\dot{x}_p - \dot{y}_p). \quad (464)$$

The right-hand side is identically zero because it is an element of the diagonal of the pushforward map, whence the conclusion follows.

The above statement, written explicitly for the unit hyper-sphere, leads indeed to non-trivial constraints on the naïve derivatives $(\dot{v}_x)_p, (\dot{w}_y)_p \in \mathbb{A}$. By using the Equation (115) for the covariant derivative, the above property reads

$$(\dot{v}_x)_p + p v_p^\top \dot{x}_p = (\dot{w}_y)_p + p w_p^\top \dot{x}_p. \quad (465)$$

Since $v_p = w_p$, the above equation may be rewritten, for example, as

$$(\dot{v}_x)_p - (\dot{w}_y)_p = p\, v_p^\top (\dot{y}_p - \dot{x}_p). \quad (466)$$

Whenever $\dot{y}_p = \dot{x}_p$, it happens trivially that $(\dot{v}_x)_p = (\dot{w}_y)_p$, otherwise, the two naïve derivatives differ by the amount $v_p^\top(\dot{y}_p - \dot{x}_p)$ along the radial direction p.

To conclude this review of the abstract properties of the principal pushforward map, let us suppose that one of the two trajectories $x(t)$ and $y(t)$ is constant, for example, suppose that

$$y(t) = y = \text{constant for every } t \in \mathbb{I}, \quad (467)$$

in such a way that its image reduces to a point-set $\{y\}$. As a consequence, we have that $\dot{y} = 0$ and therefore $\nabla_{\dot{y}}w_y = 0$ and $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,v_x}(\dot{y}) = 0$. Nevertheless, the vector field $v_x(t) = \mathbb{P}^{y\to x(t)}(w_y)$ changes over time and its covariant derivative along the velocity field $\dot{x}$ is still given by the relation (455) that, by the hypothesis made, reads:

$$(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}(\dot{x}) = \nabla_{\dot{x}}\mathbb{P}^{y\to x}(w_y)\big|_{y(t)=\text{const}}. \quad (468)$$

The above relation supplies a meaning to the principal pushforward map $(\mathrm{d}_x\mathbb{P}^{y\to x})_{x,y,w_y}$ as the covariant derivative of the transported 'constant' vector field w_y.

6.4.2. Relationship between the Principal Pushforward Map and the Curvature Endomorphism

Parallel transport does not behave trivially under composition. For instance, given three points $x, y, z \in \mathbb{M}$ not aligned along the same geodesic arc, in general it happens that $\mathbb{P}^{y\to x} \circ \mathbb{P}^{z\to y} \neq \mathbb{P}^{z\to x}$. Such fundamental observations leads to the *holonomy* theory which is closely related to the theory of curvature manifold by the *Ambrose–Singer theorem* [57]. The aim of the current section is to explore and exemplify such relationship with some details.

To start with, let us recall that to any point $x \in \mathbb{M}$ is associated an algebraic structure $\mathrm{Hol}_x(\mathbb{M})$ termed *holonomy group* formed by the set of linear operators that realize a parallel transport along any piecewise smooth closed loops originating and ending at the basepoint x. An instance of piecewise smooth closed loop easy to visualize is a *geodesic triangle* determined by three non-aligned points $x, y, z \in \mathbb{M}$ and by the three geodesic arcs that connect them. Then, given a tangent vector $w_y \in T_y\mathbb{M}$, one may start by transporting the vector w_y from the point y to the point z along the geodesic arc originating at the point y and ending in the point z by means of the parallel transport operator $\mathbb{P}^{y \to z}$, then one may transport the resulting tangent vector from the point z to the point x along the geodesic line connecting such points by means of the parallel transport operator $\mathbb{P}^{z \to x}$, and then one may transport the resulting tangent vector from the point x back to the point y along the geodesic line connecting such pair of points by means of the parallel transport operator $\mathbb{P}^{x \to y}$. The resulting tangent vector, that we denote as

$$\bar{w}_y := \left(\mathbb{P}^{x \to y} \circ \mathbb{P}^{z \to x} \circ \mathbb{P}^{y \to z} \right)(w_y) \tag{469}$$

will, in general, differ from the vector w_y that one started with, because of the curvature of the manifold. The compound linear map $\mathbb{H}_y(z, x) := \mathbb{P}^{x \to y} \circ \mathbb{P}^{z \to x} \circ \mathbb{P}^{y \to z}$ is an element of the holonomy group, namely, $\mathbb{H}_y \in \mathrm{Hol}_y(\mathbb{M})$. Notice that, since each element of the parallel transport chain in the definition of the operator $\mathbb{H}$ realizes an isometry, the map $\mathbb{H}_y$ represents an isometry as well, therefore $\|\bar{w}_y\|_y = \|w_y\|_y$.

An interesting relationship between the notion of principal pushforward map $(\mathrm{d}_x\mathbb{P}^{y \to x})$ and that of manifold holonomy descends directly from the definition (426). In fact, such definition may be rewritten equivalently as

$$
\begin{aligned}
&(\mathrm{d}_x\mathbb{P}^{y \to x})_{x(t), y(t), w_{y(t)}}(\dot{x}(t)) \\
&= \lim_{s \to 0} \tfrac{1}{s} \left\{ \left(\mathbb{P}^{x(t+s) \to x(t)} \circ \mathbb{P}^{y(t) \to x(t+s)} \right)(w_{y(t)}) - \mathbb{P}^{y(t) \to x(t)}(w_{y(t)}) \right\}, \\
&= \lim_{s \to 0} \tfrac{1}{s} \mathbb{P}^{y(t) \to x(t)} \left\{ \left(\mathbb{P}^{x(t) \to y(t)} \circ \mathbb{P}^{x(t+s) \to x(t)} \circ \mathbb{P}^{y(t) \to x(t+s)} \right)(w_{y(t)}) - w_{y(t)} \right\}, \\
&= \mathbb{P}^{y(t) \to x(t)} \left\{ \lim_{s \to 0} \tfrac{1}{s} \left(\left(\mathbb{P}^{x(t) \to y(t)} \circ \mathbb{P}^{x(t+s) \to x(t)} \circ \mathbb{P}^{y(t) \to x(t+s)} \right)(w_{y(t)}) - w_{y(t)} \right) \right\}.
\end{aligned}
\tag{470}
$$

The last line of the above relationship reveals that the argument of the limit involves a comparison between the vector w_y and the result of its transformation by the operator $\mathbb{H}_{y(t)}(x(t+s), x(t)) = \mathbb{P}^{x(t) \to y(t)} \circ \mathbb{P}^{x(t+s) \to x(t)} \circ \mathbb{P}^{y(t) \to x(t+s)}$ belonging to the holonomy group $\mathrm{Hol}_y(\mathbb{M})$ and, in particular, the relationship

$$(\mathrm{d}_x\mathbb{P}^{y \to x})_{x, y, w_y}(\dot{x}) = \mathbb{P}^{y \to x} \left\{ \lim_{s \to 0} \tfrac{1}{s} \left(\mathbb{H}_y(x(\star + s), x) - \mathrm{id}_y \right) w_y \right\} \tag{471}$$

holds true. In turn, as we have already observed, the above-defined operator $\mathbb{H}$ represents parallel transport along a geodesic triangle, which is an instance of closed loop. Since the curvature endomorphism arises from parallel transport along an infinitesimal closed loop, it might be guessed how parallel transport along a non-infinitesimal loop may be expressed as a path-integral of the curvature.

In order to clarify the relationship that relates the principal pushforward map to the curvature endomorphism, it is instrumental to recall a few results from manifold calculus.

As a first result to recall, which stands by itself as a fundamental finding in manifold calculus, let us survey the notion of *anti-covariant-derivative*, an homologous of the notion of anti-derivative in standard calculus. Such notion may be expressed in the following way: Let $\gamma : [0, 1] \to \mathbb{M}$ denote a curve and $w \in \Gamma(\mathbb{M})$ a smooth vector field defined at least on the image of γ and such that $w_{\gamma(0)} = 0$; then one might wonder how to recover the vector field w from its covariant derivative $v_{\gamma(t)} := \nabla_t^\gamma w$, supposed to be known. The answer is given by the anti-covariant-derivative formula

$$w_{\gamma(t)} = \int_0^t \mathbb{P}_\gamma^{\tau \to t}(v_{\gamma(\tau)}) d\tau. \tag{472}$$

Let us verify the correctness of such assertion. By definition of covariant derivative, it holds that

$$
\begin{aligned}
\nabla_t^\gamma w &= \lim_{s \to 0} \frac{1}{s} \left\{ \mathbb{P}_\gamma^{t+s \to t} \int_0^{t+s} \mathbb{P}_\gamma^{\tau \to t+s}(v_{\gamma(\tau)}) d\tau - \int_0^t \mathbb{P}_\gamma^{\tau \to t}(v_{\gamma(\tau)}) d\tau \right\} \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \int_0^{t+s} (\mathbb{P}_\gamma^{t+s \to t} \circ \mathbb{P}_\gamma^{\tau \to t+s})(v_{\gamma(\tau)}) d\tau - \int_0^t \mathbb{P}_\gamma^{\tau \to t}(v_{\gamma(\tau)}) d\tau \right\} \\
&= \lim_{s \to 0} \frac{1}{s} \left\{ \int_0^t (\mathbb{P}_\gamma^{t+s \to t} \circ \mathbb{P}_\gamma^{\tau \to t+s} - \mathbb{P}_\gamma^{\tau \to t})(v_{\gamma(\tau)}) d\tau + \int_t^{t+s} (\mathbb{P}_\gamma^{t+s \to t} \circ \mathbb{P}_\gamma^{\tau \to t+s})(v_{\gamma(\tau)}) d\tau \right\}.
\end{aligned}
\tag{473}
$$

Moreover, let us recall that parallel transport enjoys composition over the same curve (no matter as to whether such curve is a geodesic or a non-geodesic line), therefore $\mathbb{P}_\gamma^{t+s \to t} \circ \mathbb{P}_\gamma^{\tau \to t+s} \equiv \mathbb{P}_\gamma^{\tau \to t}$. As a consequence, the first integral in the last line of the above multiline equation is identically zero, while the second integral equals

$$\int_t^{t+s} \mathbb{P}_\gamma^{\tau \to t}(v_{\gamma(\tau)}) d\tau = s\, v_{\gamma(t)} + \mathcal{O}(s^2), \tag{474}$$

by a continuity argument, hence the assertion (472) follows.

The second result that is convenient to recall is based on the notion of *geodesic homotopy*. Let us define the function $h : [-\epsilon,\, \epsilon] \times [0,\, 1] \to \mathbb{M}$ to be a geodesic homotopy, namely, a family of curves such that for any fixed $s \in [-\epsilon,\, \epsilon]$, $t \mapsto h(s,t)$ traces a geodesic curve on $\mathbb{M}$ connecting the point $h(s,0)$ to the point $h(s,1)$. The image of an homotopy is sometimes referred to as *homotopic net*. In addition, let us define $\sigma : [-\epsilon,\, \epsilon] \times [0,\, 1] \to T\mathbb{M}$ to be a smooth vector field parallel along the homotopy, namely, such that $\sigma(s,t) \in T_{h(s,t)}\mathbb{M}$ and

$$(\nabla_{\partial_s h}\sigma)(s,0) = 0 \text{ and } \nabla_{\partial_t h}\sigma(s,t) = 0 \tag{475}$$

for every value of the index $s \in [-\epsilon,\, \epsilon]$ and for every value of the affine parameter $t \in [0,\, 1]$. Namely, the vector field σ is parallel along the back-end curve $h(s,0)$ and along each geodesic $h(\star,t)$ of the homotopic net. Notice that one such vector field is entirely determined by its value $\sigma(0,0)$ given and by the homotopy itself. Then, the following relationship holds:

$$(\nabla_{\partial_s h}\sigma)(s,1) = \left(\int_0^1 \mathbb{P}^{h(s,t) \to h(s,1)} \circ \mathbf{R}_{h(s,t)}^{\partial_t h, \partial_s h} \circ \mathbb{P}^{h(s,1) \to h(s,t)} dt \right) \sigma(s,1), \tag{476}$$

where $\mathbf{R}$ denotes the Riemannian curvature endomorphism associated to the manifold and to its Levi-Civita connection. Namely, the transversal covariant derivative of the vector field along the front-end curve $h(s,1)$ is given by a linear operator, which depends on the curvature of the manifold along the homotopic net, applied to the field. In the following, to ease notation we shall denote as $\xi := \partial_s h$ the transversal velocity and as $v := \partial_t h$ the velocity along each geodesic belonging to the homotopic net. By definition of Riemannian curvature endomorphism, we have that

$$\nabla_v \nabla_\xi \sigma = \nabla_\xi \nabla_v \sigma + \nabla_{[v,\xi]}\sigma + \mathbf{R}_h^{v,\xi}\sigma. \tag{477}$$

As we had already observed earlier, the vector fields v and ξ commute, moreover, by hypothesis, the vector field σ is parallel along the homotopy, therefore $\nabla_v \sigma = 0$. As a consequence, the expression (477) simplifies to

$$(\nabla_v \nabla_\xi \sigma)(s,t) = \mathbf{R}_{h(s,t)}^{v(s,t),\xi(s,t)}\sigma(s,t). \tag{478}$$

By the assumption of parallelism, for any given index s, the vector $\sigma(s,t)$ may be recovered from its front-end value $\sigma(s,1)$ through parallel transport, hence

$$(\nabla_v \nabla_{\xi} \sigma)(s,t) = \mathbf{R}^{v(s,t),\xi(s,t)}_{h(s,t)} \mathbb{P}^{h(s,1)\to h(s,t)} \sigma(s,1). \tag{479}$$

Applying the notion of anti-covariant-derivative to the left-hand side of the expression (479) gives

$$(\nabla_{\xi} \sigma)(s,t) = \int_0^t \mathbb{P}^{h(s,\tau)\to h(s,t)} \left(\mathbf{R}^{v(s,\tau),\xi(s,\tau)}_{h(s,\tau)} \mathbb{P}^{h(s,1)\to h(s,\tau)} \sigma(s,1) \right) d\tau. \tag{480}$$

Setting $t = 1$ gives the sought result.

On the basis of the Equation (476), it is possible to express the principal pushforward map in terms of the Riemannian curvature endomorphism. Such relationship may be made explicit by observing that the principal pushforward map may be expressed in closed form as

$$\begin{cases} (d_x \mathbb{P}^{y\to x})_{x,y,w_y}(\dot{x}) = (\mathbf{F}^0_{x,y} \circ \mathbb{P}^{y\to x})(w_y), \text{ where} \\ \mathbf{F}^0_{x,y} := \int_0^1 \mathbb{P}^{\exp_y(t\log_y x)\to x} \circ \mathbf{R}^0_{x,y,\dot{x}}(t) \circ \mathbb{P}^{x\to \exp_y(t\log_y x)} t dt, \text{ with} \\ \mathbf{R}^0_{x,y,\dot{x}}(t) := \mathbf{R}_{\exp_y(t\log_y x)}((d\exp)_{y,t\log_y x}(\log_y x), (d\exp)_{y,t\log_y x}((d_x \log)_{y,x}(\dot{x}))). \end{cases} \tag{481}$$

Such statement may be explained by choosing an appropriate homotopy h and a vector field σ. In particular, let us choose the homotopy $h : [-\epsilon, \epsilon] \times [0, 1] \to \mathbb{M}$ as

$$h(s,t) := \exp_y(t\log_y \exp_x(s\dot{x})), \tag{482}$$

where the back-end curve $s \mapsto h(s,0)$ is such that $h(0,0) = y \in \mathbb{M}$ and the front-end curve $s \mapsto h(s,1)$ is such that $\partial_s h(s,1)|_{s=0} = \dot{x}(s)|_{s=0}$. In addition, we have denoted, for simplicity, by $\dot{x}$ the tangent vector $\dot{x}(0)$. The vector field $\sigma : [-\epsilon, \epsilon] \times [0, 1] \to T_{h(s,t)}\mathbb{M}$ along the corresponding homotopic net is defined as

$$\sigma(s,t) := \mathbb{P}^{y\to h(s,t)}(w_y), \tag{483}$$

with $w_y \in T_y\mathbb{M}$ fixed. Apparently, the homotopy (482) traces a smooth geodesic net, in fact, for s fixed, $t \mapsto h(s,t)$ traces a geodesic line that departs from the point y and ends at the point $\exp_x(s\dot{x})$. Moreover, the vector field (483) satisfies both conditions (475), in fact, for any fixed value of the index s, the vector field $t \mapsto \sigma(\star,t)$ arises as the result of parallel translating the vector w_y along the corresponding geodesic of the homotopy net, hence $\sigma(\star,t)$ is parallel along each geodesic; in addition, the tangent value $\sigma(s,0)$ is obtained by parallel transporting the vector w_y along the curve $h(s,0)$, hence the vector field $s \mapsto \sigma(s,0)$ is parallel along the back-end arc of which ξ represents the velocity. Therefore, from the relationship (476), it follows that

$$\begin{cases} (\nabla_{\xi} \sigma)(s,1) = \mathbf{F}(s)(\sigma(s,1)), \text{ with} \\ \mathbf{F}(s) := \int_0^1 \mathbb{P}^{h(s,t)\to h(s,1)} \circ \mathbf{R}^{v(s,t),\xi(s,t)}_{h(s,t)} \circ \mathbb{P}^{h(s,1)\to h(s,t)} dt, \end{cases} \tag{484}$$

where $\mathbf{F}(s) \in \text{End}(T_{h(s,1)}\mathbb{M})$ for every $s \in [-\epsilon, \epsilon]$. The relationship between the covariant derivative $(\nabla_{\xi} \sigma)$ and the principal pushforward map $(d_x \mathbb{P}^{y\to x})_{x,y,w_y}(\dot{x})$ is given by the observation that

$$(d_x \mathbb{P}^{y\to x})_{h(s,1),y,w_y}(\xi(s,1)) = \nabla^{h(s,1)}_s \mathbb{P}^{y\to x(s)}(w_y)$$
$$= (\nabla_{\xi} \sigma)(s,1). \tag{485}$$

Therefore, on the fiducial geodesic corresponding to the value $s = 0$, it holds that

$$\begin{cases} (\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w_y}(\dot{x}) = \mathbf{F}_0(\sigma(0,1)), \text{ with} \\ \mathbf{F}_0 = \int_0^1 \mathbb{P}^{h(0,t) \to h(0,1)} \circ \mathbf{R}_{h(0,t)}^{v(0,t),\xi(0,t)} \circ \mathbb{P}^{h(0,1) \to h(0,t)} \mathrm{d}t, \end{cases} \tag{486}$$

where, in particular, $h(0,t) = \exp_y(t \log_y x)$. The partial derivative of the homotopy (482) with respect to the parameter t, namely the velocity along each geodesic of the homotopic net, reads

$$(\partial_t h)(s,t) = (\mathrm{d}_x \exp)_{y, t \log_y \exp_x(s\dot{x})}(\log_y \exp_x(s\dot{x})), \tag{487}$$

where $\mathrm{d}_x \exp : (T\mathbb{M})^2 \to T\mathbb{M}$ denotes the pushforward map of the exponential map $(x,v) \mapsto \exp_x(v)$ with respect to the x argument, therefore the velocity of the fiducial geodesic reads

$$v(0,t) = (\mathrm{d} \exp)_{y, t \log_y x}(\log_y x). \tag{488}$$

Instead, the partial derivative of the homotopy (482) computed with respect to the index s reads

$$(\partial_s h)(s,t) = (\mathrm{d}_x \exp)_{y, t \log_y \exp_x(s\dot{x})}\Big(t(\mathrm{d}_x \log)_{y, \exp_x(s\dot{x})}(\mathrm{d}_x \exp)_{x, s\dot{x}}(\dot{x})\Big), \tag{489}$$

where $\mathrm{d}_x \log : \mathbb{M}^2 \times T\mathbb{M} \to T\mathbb{M}$ denotes the pushforward map of the logarithmic map $(y,x) \mapsto \log_y x$ with respect to the x argument. Such partial derivative at $s = 0$ reads

$$\xi(0,t) = (\mathrm{d}_x \exp)_{y, t \log_y x}\big(t(\mathrm{d}_x \log)_{y,x}(\mathrm{d} \exp)_{x,0}(\dot{x})\big). \tag{490}$$

Since $(\mathrm{d} \exp)_{x,0} = \mathrm{id}_x$ and the pushforward map $d \exp$ is linear, the above relation may be simplified to

$$\xi(0,t) = t((\mathrm{d}_x \exp)_{y, t \log_y x} \circ (\mathrm{d}_x \log)_{y,x})(\dot{x}). \tag{491}$$

Such expression emphasizes the fact that the transversal velocity $\xi(0,t)$ is linear in the vector $\dot{x}$. In conclusion, replacing the Equations (488) and (491) into the expression (486), we get:

$$\begin{cases} (\mathrm{d}_x \mathbb{P}^{y \to x})_{x,y,w_y}(\dot{x}) = (\mathbf{F}_0 \circ \mathbb{P}^{y \to x})(w_y), \\ \mathbf{F}_0 = \int_0^1 \mathbb{P}^{h(0,t) \to h(0,1)} \circ \mathbf{R}_{h(0,t)}^{v(0,t),\xi(0,t)} \circ \mathbb{P}^{h(0,1) \to h(0,t)} \mathrm{d}t, \\ v(0,t) = (\mathrm{d} \exp)_{y, t \log_y x}(\log_y x), \\ \xi(0,t) = t((\mathrm{d}_x \exp)_{y, t \log_y x} \circ (\mathrm{d}_x \log)_{y,x})(\dot{x}), \end{cases} \tag{492}$$

which justifies the assertion (481). Notice that the integral in the second equation of (481) is well defined because the integrand is a linear endomorphism in $\mathrm{End}(T_x\mathbb{M})$.

Let us apply the above formula to the unit hypersphere as an example of calculation.

Example 33. *For the unit hypersphere $\mathbb{S}^{n-1}$ endowed with the canonical metric we have determined that the curvature endomorphism reads*

$$\mathbf{R}^{u,v}(w) := (v^\top w)u - (u^\top w)v, \tag{493}$$

for every $u, v, w \in T_x\mathbb{S}^{n-1}$ and $x \in \mathbb{S}^{n-1}$. Since the curvature endomorphism does not depend explicitly on the point x, we omitted its indication.

A matrix representation of the curvature endomorphism $\mathbf{R}$ is the curvature tensor $[\![\mathbf{R}^{u,v}]\!] := uv^\top - vu^\top$, from which the bilinearity and antisymmetry properties are easily verified. For the ordinary sphere $\mathbb{S}^2$, denoting by $\wedge : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ the outer product in $\mathbb{R}^3$, the skew-symmetric matrix $[\![\mathbf{R}^{u,v}]\!]$ coincides with the matrix representation of the 3-vector $u \wedge v$, namely $[\![\mathbf{R}^{u,v}]\!]^\vee = u \wedge v$.

Let us recall that the exponential map is expressed as

$$\exp_y(w) = \begin{cases} y\,\cos(\|w\|) + w\,\text{sinc}(\|w\|), & \text{if } w \neq 0, \\ y, & \text{if } w = 0, \end{cases} \tag{494}$$

therefore, its inverse, the logarithmic map, is expressed by the relation:

$$\log_y x = \frac{(I_n - yy^\top)x}{\text{sinc}\,\mathscr{D}(x,y)}, \tag{495}$$

where it is assumed that the points $x, y \in \mathbb{S}^{n-1}$ are not antipodal. In the above formula, the symbol sinc : $\mathbb{R} \to \mathbb{R}$ denotes the cardinal sine function defined as:

$$\text{sinc}\,\zeta := \begin{cases} \zeta^{-1}\sin\zeta & \text{for } \zeta \neq 0, \\ 1 & \text{for } \zeta = 0, \end{cases} \tag{496}$$

and the geodesic distance between two points x, $y \in \mathbb{S}^{n-1}$ associated to the canonical metrics reads

$$\mathscr{D}(x,y) = \|\log_y x\|_y = \text{acos}(x^\top y), \tag{497}$$

where the inverse cosine function ranges in $[0, \pi]$.

   Matrix representations of the derivatives of the exponential and the logarithmic maps are computed to be

$$[\![(\text{d}\exp)_{y,w}]\!] = \cos(\|w\|)I_n - \text{sinc}(\|w\|)yw^\top, \tag{498}$$

$$[\![(\text{d}_x\log)_{y,x}]\!] = \frac{I_n - yy^\top}{\text{sinc}^2\mathscr{D}(x,y)}\big[\,\text{sinc}\,\mathscr{D}(x,y)I_n - (\,\text{sinc}'\,\mathscr{D}(x,y))\,\text{acos}'(x^\top y)xy^\top\,\big], \tag{499}$$

where the prime denotes derivation. The corresponding matrix representation of the endomorphism $\mathbf{R}^0$ takes the form:

$$[\![\mathbf{R}^0]\!] = [\![(\text{d}\exp)_{y,t\log_y x}]\!]\Big\{(\log_y x)\dot{x}^\top[\![(\text{d}_x\log)_{y,x}]\!]^\top - [\![(\text{d}_x\log)_{y,x}]\!]\dot{x}\log_y^\top x\Big\}[\![(\text{d}\exp)_{y,t\log_y x}]\!]^\top, \tag{500}$$

 which denotes an anti-symmetric matrix.

7. Discrete-Time Dynamical Systems on Manifolds

   Dynamical systems on manifolds may appear under the form of *discrete-time* systems as well as *continuous-time systems*. In turn, discrete-time systems may be *natively discrete-time* or may be derived from continuous-time systems through time sampling. Discrete-time dynamical systems are nowadays conceived as algorithms and implemented as pieces of codes running on computing platforms.

   The aim of the present last section of this tutorial paper is to recall an *auto-regressive, moving average (ARMA) mathematical model* from [58] that generates a discrete-time temporal sequence on a smooth manifold. In such setting, not only the algorithmic structure of the ARMA model reflect the mathematical structure of the state manifold but the input and the output of the ARMA dynamical system are structured sequences that belong to the tangent bundle $T\mathbb{M}$ of same smooth manifold $\mathbb{M}$.

   The present section is organized in the following way. The Section 7.1 describes second-order discrete-time auto-regressive moving-average dynamical systems on Riemannian manifolds. To end with, Section 7.2 illustrates examples of ARMA-type dynamical systems on manifold.

7.1. Auto-Regressive Moving-Average (ARMA) Systems on Riemannian Manifolds

   An ARMA model on a manifold is described in terms of a pair of sequences $(x_k, v_k) \in T\mathbb{M}$, where x_k denotes the output signal of the model, v_k denotes the state of the model

and $k \in \mathbb{Z}$ denotes a discrete-time index. In the familiar case that $\mathbb{M} \equiv \mathbb{R}^p$, choosing the standard identification $T_x \mathbb{R}^p \cong \mathbb{R}^p$, for any $x \in \mathbb{R}^p$, an ARMA(r, q) model on the tangent bundle $T\mathbb{R}^p$ appears under the form of a state-evolution equation and of a state-to-output transformation equation. In particular, the equation that governs the evolution of the internal state $v_k \in T_{x_k} \mathbb{R}^p$ reads:

$$v_k = \underbrace{\sum_{i=1}^{r} A_i v_{k-i}}_{\substack{\text{Auto-regressive} \\ \text{(AR)}}} + \underbrace{\sum_{i=0}^{q} B_i w_{k-i}}_{\substack{\text{Moving average} \\ \text{(MA)}}}, \tag{501}$$

where the *transition matrices* $A_i \in \mathbb{R}^{p \times p}$ and $B_i \in \mathbb{R}^{p \times p}$ denote the parameters of the auto-regressive and of the moving-average subsystems, respectively, and the sequence $w_k \in T_{x_k} \mathbb{R}^p$ denotes a white noise. Let us recall the salient features of the two subsystem of an ARMA model:

- **Moving average (MA) subsystem**: The moving-average subsystem consists of a linear combination of the current value and of previous samples of the *input* signal;
- **Auto-regressive (AR) subsystem**: The auto-regressive subsystem consists of a linear combination of previous values of the internal *state* of the system.

The integer $r \geqslant 1$ denotes the memory depth of the AR subsystem, while the integer $q \geqslant 0$ denotes the memory depth of the MA subsystem. (Notice that speaking of 'memory depth' for the AR subsystem is purely conventional since an AR subsystem has an infinite memory span.) In addition, the equation that describes how the internal state v_k transforms into the output $x_k \in \mathbb{R}^p$ reads:

$$x_{k+1} = x_k + v_k + c, \tag{502}$$

with $c \in \mathbb{R}^p$ denoting a constant bias (or *drift* term). It is worth remarking that, since the tangent bundle $T\mathbb{R}^p$ is trivial (in fact, $T\mathbb{R}^p$ is isomorphic to $\mathbb{R}^p \times \mathbb{R}^p$), the Equation (501) is based on vector addition. On a general Riemannian manifold $\mathbb{M}$, however, vector addition in the tangent bundle $T\mathbb{M}$ (that is, in the AR part and in the MA part) may be hardly defined without invoking the concept of parallel transport. Likewise, Equation (502) is based on vector addition on $\mathbb{R}^p$, which is not defined on a general (curved) Riemannian manifold and needs rather to be replaced by the exponential map.

The proposed extension of the above ARMA system to a Riemannian manifold $\mathbb{M}$ is based on the following considerations:

- At each step k, it is necessary to compute a linear combination of several tangent vectors in $T_{x_k} \mathbb{M}$ via the transition operators $\mathcal{A}_i : (x, v) \in T\mathbb{M} \mapsto w \in T_x \mathbb{M}$ for the auto-regressive subsystem and $\mathcal{B}_i : (x, v) \in T\mathbb{M} \mapsto w \in T_x \mathbb{M}$ for the moving-average subsystem, where the operators $\mathcal{A}_i$ and $\mathcal{B}_i$ are linear in the second (namely, the vectorial) argument.
- In order to compute the linear combination of previous and current values of the state variable, at each step k it is necessary to parallel transport each tangent vector v_{k-1}, ..., v_{k-r}, w_{k-1}, ..., w_{k-q} to the tangent space $T_{x_k} \mathbb{M}$.
- In order to propagate the constant bias, at each step k it is necessary to parallel transport it to the tangent space $T_{x_k} \mathbb{M}$.

On the basis of the above considerations, a manifold-type auto-regressive, moving average M-ARMA$(2, q)$ dynamical system on the tangent bundle $T\mathbb{M}$ of the manifold $\mathbb{M}$ may be cast as:

$$\begin{cases} v_k = \sum_{i=1}^{2} \mathcal{A}_i(x_k, \mathbb{P}^{x_{k-i} \to x_k}(v_{k-i})) + \sum_{i=0}^{q} \mathcal{B}_i(x_k, \mathbb{P}^{x_{k-i} \to x_k}(w_{k-i})), \\ x_{k+1} = \exp_{x_k}(v_k + \mathbb{P}^{x_0 \to x_k}(c)), \end{cases} \tag{503}$$

for $k = 0, 1, 2, \ldots$. At $k = 0$, it is necessary to know the initial seed $x_0 \in \mathbb{M}$ and to put into effect an extension of standard zero-padding to account for lagged vector-samples that are not available at $k = 0$:

$$
\begin{cases}
v_{-1} = v_{-2} = 0, \\
w_{-1} = w_{-2} = \ldots = w_{-q} = 0, \\
x_{-1} = x_{-2} = \ldots = x_{-\beta} = x_0,
\end{cases}
\tag{504}
$$

with $\beta := \max\{2, q\}$ denoting the memory depth of the whole M-ARMA system. Notice that the memory depth on the state variable v has been chosen to be 2 since this tutorial paper cover mostly second-order systems and for illustrative purposes. The tangent state-vectors v_k are computed as the output of an auto-regressive moving-average system on the tangent bundle $T\mathbb{M}$. In a M-ARMA(r, q) system whose dynamics manifests through a state-sequence $v_k \in T_{x_k}\mathbb{M}$ as a response to an input sequence $w_k \in T_{x_k}\mathbb{M}$, the current state variable value v_k depends from the past state-variable values v_{k-1} and v_{k-2} and from the current and past input-variable values $w_k, w_{k-1}, \ldots, w_{k-q}$. The additional term $\mathbb{P}^{x_0 \to x_k}(c)$ in the second equation of the mathematical model (503) represents a drift term that causes a slight change in the current state even when the scalar velocity $\|v_k\|_{x_k}$ is close to zero.

The M-ARMA model (503), may be cast in a different way by replacing parallel transport with vector transport. The corresponding alternative version of such M-ARMA system reads

$$
\begin{cases}
v_k = \sum_{i=1}^{2} \mathcal{A}_i(x_k, \Pi_{x_k}(v_{k-i})) + \sum_{i=0}^{q} \mathcal{B}_i(x_k, \Pi_{x_k}(w_{k-i})), \\
x_{k+1} = \exp_{x_k}(v_k + \Pi_{x_k}(c)),
\end{cases}
\tag{505}
$$

for $k = 0, 1, 2, \ldots$. In some occasions, even if a parallel transport formula is available, vector transport is invoked due to its alleged lighter computational burden.

The steps required to implement the M-ARMA$(2, q)$ models (503) and (505) are summarized in Algorithm 4.

Algorithm 4 Pseudocode to implement the M-ARMA$(2, q)$ systems (503) and (505).

1: Choose a metric for the manifold $\mathbb{M}$ and make sure the exponential map $\exp : T\mathbb{M} \to \mathbb{M}$ and the parallel transport operator $\mathbb{P}^{x \to y} : T_x\mathbb{M} \to T_y\mathbb{M}$ or the vector transport operator π_x, with $x, y \in \mathbb{M}$, are available

2: Choose the initial seed $x_0 \in \mathbb{M}$ and the drift constant $c \in T_{x_0}\mathbb{M}$

3: Choose 2 transition operators $\mathcal{A}_i : T\mathbb{M} \to T\mathbb{M}$ and q transition operators $\mathcal{B}_i : T\mathbb{M} \to T\mathbb{M}$

4: Set $\beta := \max\{2, q\}$ and set up the zero-padding $v_{-1} = v_{-2} = 0$, $w_{-1} = w_{-2} = \ldots = w_{-q} = 0$, $x_{-1} = x_{-2} = \ldots = x_{-\beta} = x_0$

5: **for** $k = 0$ to K **do**

6:　　Get a tangent vector $w_k \in T_{x_k}\mathbb{M}$ from an input source

7:　　Compute tangent vectors $\tilde{v}_{k,i} := \mathbb{P}^{x_{k-i} \to x_k}(v_{k-i})$ and $\tilde{w}_{k,j} := \mathbb{P}^{x_{k-j} \to x_k}(w_{k-j})$, or $\tilde{v}_{k,i} := \Pi_{x_k}(v_{k-i})$ and $\tilde{w}_{k,j} := \Pi_{x_k}(w_{k-j})$, for $i = 1, 2$ and $j = 0, \ldots, q$ and the actualized drift term $d_k = \mathbb{P}^{x_0 \to x_k}(c)$ or $d_k = \Pi_{x_k}(c)$

8:　　Compute the next state-velocity term $v_k = \sum_{i=1}^{2} \mathcal{A}_i(x_k, \tilde{v}_{k,i}) + \sum_{j=0}^{q} \mathcal{B}_j(x_k, \tilde{v}_{k,j})$

9:　　Compute the next term in the output sequence $x_{k+1} = \exp_{x_k}(v_k + d_k)$

10: **end for**

By adding extra terms to the right-hand side of the first equation of the system (503) or (505), the discussed M-ARMA model may be extended to a M-ARMAX (M-ARMA with exogenous inputs, see, for instance [59]).

7.2. Examples of M-ARMA-Type Dynamical Systems

The aim of the present section is to illustrate some examples of the discussed M-ARMA model. The general M-ARMA$(2, q)$ system and the notion of autocorrelation function on manifold will be illustrated through examples on the unit hypersphere, the manifold

of symmetric positive-definite matrices and the compact Stiefel manifold. In particular, examples about how to choose the auto-regressive and the moving-average transition operators $\mathcal{A}_i$ and $\mathcal{B}_i$ will be given. In particular, in the following we shall assume that the purpose of running a M-ARMA algorithm is to generate a non-white pseudo-random sequence on a given manifold, on the basis of a random number generator that is used to produce a pseudo-random input sequence w_k.

7.2.1. Generation of Pseudo-Random Paths on the Unit Hyper-Sphere

In order to generate a pseudo-random vector $v \in T_x\mathbb{S}^{p-1}$, given a point $x \in \mathbb{S}^{p-1}$, the following algorithm may be employed:

- Generate a random vector $\tilde{v} \in \mathbb{R}^p$ by stacking the outputs of a pseudo-random scalar generator;
- Project the vector $\tilde{v}$ into the tangent space $T_x\mathbb{S}^{p-1}$ by $v = \Pi_x\{\tilde{v}\}$.

The generation of a pseudo-random path over the hyper-sphere $\mathbb{S}^{p-1}$ requires to choose the transition operators that take part in the M-ARMA$(2, q)$ model (503). In particular, a possible choice for the transition operators $\mathcal{A}_i : (x, v) \in T\mathbb{S}^{p-1} \mapsto w \in T_x\mathbb{S}^{p-1}$ and $\mathcal{B}_i : (x, v) \in T\mathbb{S}^{p-1} \mapsto w \in T_x\mathbb{S}^{p-1}$ is:

$$\mathcal{A}_i(x, v) = \Pi_x\{A_i v\}, \ i = 1, \ 2, \ \mathcal{B}_i(x, v) = \Pi_x\{B_i v\}, \ i \geqslant 0, \tag{506}$$

where the square matrices $A_i, B_i \in \mathbb{R}^{p \times p}$ may be chosen arbitrarily.

Example 34. *To give an example of the above setting, let us choose $p = 3$ and $q = 4$. Moreover, let us opt for the version based on parallel transport. The resulting pseudo-random path-generation algorithm M-ARMA$(2, 4)$ on the hypersphere $\mathbb{S}^2$ may be particularized from the general expression (503) as follows:*

$$
\begin{aligned}
v_k &= (I_3 - x_k x_k^\top)A_1\mathbb{P}^{x_{k-1} \to x_k}(v_{k-1}) + (I_3 - x_k x_k^\top)A_2\mathbb{P}^{x_{k-2} \to x_k}(v_{k-2}) + (I_3 - x_k x_k^\top)B_0 w_k + \\
&= (I_3 - x_k x_k^\top)B_1\mathbb{P}^{x_{k-1} \to x_k}(w_{k-1}) + (I_3 - x_k x_k^\top)B_2\mathbb{P}^{x_{k-1} \to x_k}(w_{k-2}) + \\
&= (I_3 - x_k x_k^\top)B_3\mathbb{P}^{x_{k-3} \to x_k}(w_{k-3}) + (I_3 - x_k x_k^\top)B_4\mathbb{P}^{x_{k-4} \to x_k}(w_{k-4}), \\
x_{k+1} &= \exp_{x_k}(v_k + \mathbb{P}^{x_0 \to x_k}(c)),
\end{aligned}
\tag{507}
$$

with $x_k \in \mathbb{S}^2$, w_k, $v_k \in T_{x_k}\mathbb{S}^2$ for $k = 0, 1, 2, \ldots, K$ and $c \in T_{x_0}\mathbb{S}^2$ generated randomly by projection over the tangent space $T_{x_0}\mathbb{S}^2$. The value of the integer K may be chosen as large as one needs.

7.2.2. Generation of Pseudo-Random Paths on the Manifold of Symmetric, Positive-Definite Matrices

The starting point is to generate a pseudo-random tangent matrix $V \in T_P\mathbb{S}^+(n)$, given a symmetric, positive-definite matrix $P \in \mathbb{S}^+(n)$. Such result may be achieved by putting into effect the following scheme:

- Generate a pseudo-random matrix $\tilde{V} \in \mathbb{R}^{n \times n}$ by means of a pseudo-random scalar generator;
- Project the matrix $\tilde{V}$ into the tangent space $T_P\mathbb{S}^+(n)$ by the rule $V = \Pi_P(\tilde{V}) = \frac{1}{2}(\tilde{V} + \tilde{V}^\top)$.

In addition, we may recognize that a way to select the transition operators $\mathcal{A}_i : (P, V) \in T\mathbb{S}^+(n) \mapsto W \in T_P\mathbb{S}^+(n)$ and $\mathcal{B}_i : (P, V) \in T\mathbb{S}^+(n) \mapsto W \in T_P\mathbb{S}^+(n)$ is:

$$
\begin{cases}
\mathcal{A}_i(P, V) = a_i(PV + VP), \ i = 1, \ 2 \\
\mathcal{B}_i(P, V) = b_i(PV + VP), \ i \geqslant 0,
\end{cases}
\tag{508}
$$

where $a_i \in \mathbb{R}$ and $b_i \in \mathbb{R}$ denote constant parameters chosen arbitrarily.

7.2.3. Generation of Pseudo-Random Paths on the Compact Stiefel Manifold

The canonical metric is the metric of choice in the present section.

With the aim to generate a pseudo-random tangent matrix $V \in T_X \mathrm{St}(n, p)$, given a point $X \in \mathrm{St}(n, p)$, the following scheme may be used:

- Generate a pseudo-random matrix $\tilde{V} \in \mathbb{R}^{n \times p}$ by a pseudo-random scalar generator;
- Project the matrix $\tilde{V}$ into the tangent space $T_X \mathrm{St}(n, p)$ by the rule $V = \tilde{V} - X \tilde{V}^\top X$.

The generation of a pseudo-random path over the manifold $\mathrm{St}(n, p)$ requires to select transition operators $\mathcal{A}_i : (X, V) \in T\mathrm{St}(n, p) \mapsto W \in T_X \mathrm{St}(n, p)$ and $\mathcal{B}_i : (X, V) \in T\mathrm{St}(n, p) \mapsto W \in T_X \mathrm{St}(n, p)$. A possible way to effect such choice is:

$$\mathcal{A}_i(X, V) = a_i V, \ i = 1, \ 2, \ \mathcal{B}_i(X, V) = b_i X V^\top X, \ i \geqslant 0, \tag{509}$$

where $a_i \in \mathbb{R}$ and $b_i \in \mathbb{R}$ denote arbitrarily chosen constant parameters.

Example 35. *For exemplification purposes, one may select $n = 5$, $p = 2$ and $q = 4$. The initial seed needs to be chosen randomly in $\mathrm{St}(5, 2)$ and one may choose, for simplicity, not to include any drift term in the equations. The pseudo-random path generation rule (505) particularizes, in this case, to:*

$$V_k = \Pi_{X_k} \left(a_1 V_{k-1} + a_2 V_{k-2} + X_k^\top (\sum_{i=0}^{4} b_i W_{k-i}) X_k \right),$$

$$X_{k+1} = \exp_{X_k}(V_k), \tag{510}$$

for $k = 0, \ 1, \ 2, \ \ldots, \ K$. The exponential is implemented as:

$$\exp_X(V) = [X \ Q] \, \mathrm{Exp} \left(\begin{bmatrix} X^\top V & -R^\top \\ R & 0_2 \end{bmatrix} \right) \begin{bmatrix} I_2 \\ 0_2 \end{bmatrix}, \tag{511}$$

where Q and R denote the factors of the compact QR decomposition of the matrix $(I_5 - XX^\top)V$.

8. Conclusions

The present tutorial paper, the second of a series of tutorials on manifold calculus with application to non-linear control systems, was devoted to topics from manifold calculus apt to describe properties of non-linear dynamical systems subject to holonomic constraints in its state variable as well as to control non-linear dynamical systems whose state belong to curved manifolds. As an interesting cases of study for illustrative purpose, special emphasis was put on designing algorithms to achieve synchronization of manifold-type systems based on feedback control and on developing numerical methods tailored to curved manifolds to implement such systems and algorithms.

Aside of practical applications to system theory and control, the present tutorial focuses even on purely theoretical topics such as curvature and covariant derivation. The present tutorial paper does not cover further topic of great interest, such as integration on manifold and systems of order larger than two. Such important arguments will be treated in a forthcoming part of this multi-part tutorial.

References

1. Bloch, A. An introduction to aspects of geometric control theory. In *Nonholonomic Mechanics and Control*; Krishnaprasad, P., Murray, R., Eds.; Interdisciplinary Applied Mathematics; Springer: New York, NY, USA, 2015; Volume 24.
2. Bullo, F.; Lewis, A. Geometric control of mechanical systems. In *Texts in Applied Mathematics*; Springer: New York, NY, USA; Heidelberg/Berlin, Germany, 2004; Volume 49.
3. Návrat, A.; Vašík, P. On geometric control models of a robotic snake. *Note Mat.* **2017**, *37*, 120–129.
4. Nijmeijer, H.; van der Schaft, A. *Nonlinear Dynamical Control Systems*; Springer: New York, NY, USA, 1990.
5. Sreenath, K.; Lee, T.; Kumar, V. Geometric control and differential flatness of a quadrotor UAV with a cable-suspended load. In Proceedings of the 52nd IEEE Conference on Decision and Control, Firenze, Italy, 10–13 December 2013; pp. 2269–2274. [CrossRef]
6. Agrachev, A.; Sachkov, Y. *Control Theory from the Geometric Viewpoint*; Springer: Berlin/Heidelberg, Germnay, 2004; Volume 87.
7. Fiori, S. Extension of PID regulators to dynamical systems on smooth manifolds (M-PID). *SIAM J. Control Optim.* **2021**, *59*, 78–102. [CrossRef]
8. Brockett, R. Lie algebras and lie groups in control theory. In *Geometric Methods in System Theory*; Mayne, D., Brockett, R., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 1973; pp. 43–82.
9. Jakubzyk, B. Introduction to geometric nonlinear control; controllability and Lie bracket. In *Mathematical Control Theory, Lectures Notes of a Minicourse*; Polish Academy of Sciences: Warsaw, Poland, 2002; pp. 107–168.
10. Jurdjevic, V. Optimal control on Lie groups and integrable Hamiltonian systems. *Regul. Chaotic Dyn.* **2011**, *16*, 514–535. [CrossRef]
11. Sastry, S. Geometric nonlinear control. In *Nonlinear Systems—Analysis, Stability, and Control*; Sastry, S., Ed.; Springer Science+Business Media: New York, NY, USA, 1999; pp. 510–573.
12. Fiori, S. Non-delayed synchronization of non-autonomous dynamical systems on Riemannian manifolds and its applications. *Nonlinear Dyn.* **2018**, *94*, 3077–3100. [CrossRef]
13. Fiori, S. Manifold calculus in system theory and control—Fundamentals and first-order systems. *Symmetry* **2021**, *13*, 2092. [CrossRef]
14. Poisson, E.; Pound, A.; Vega, I. The motion of point particles in curved spacetime. *Living Rev. Relativ.* **2011**, *14*, 7 . [CrossRef]
15. Ferreira, R.; Xavier, J. Hessian of the Riemannian squared distance function on connected locally symmetric spaces with applications. In Proceedings of the 7th Portuguese conference on automatic control (Controlo 2006), Lisbon, Portugal, 11–13 September 2006.
16. Pennec, X. *Hessian of the Riemannian Squared Distance*; Technical report; Université Côte d'Azur and Inria Sophia-Antipolis Méditerranée: Valbonne, France, 2017.
17. Fiori, S. Gyroscopic signals smoothness assessment by geometric jolt estimation. *Math. Methods Appl. Sci.* **2017**, *40*, 5893–5905. [CrossRef]
18. Fiori, S. Error-based control systems on Riemannian state manifolds: Properties of the principal pushforward map associated to parallel transport. *Math. Control Relat. Fields* **2021**, *11*, 143–167. [CrossRef]
19. Röbenack, K. Computation of multiple Lie derivatives by algorithmic differentiation. *J. Comput. Appl. Math.* **2008**, *213*, 454–464. [CrossRef]
20. Civita, A.; Fiori, S.; Romani, G. A mobile acquisition system and a method for hips sway fluency assessment. *Information* **2018**, *9*, 321. [CrossRef]
21. Fiori, S. A closed-form expression of the instantaneous rotational lurch index to evaluate its numerical approximation. *Symmetry* **2019**, *11*, 1208. [CrossRef]
22. Ferreira, R.; Xavier, J.; Costeira, J.; Barroso, V. Newton method for Riemannian centroid computation in naturally reductive homogeneous spaces. In Proceedings of the 2006 IEEE International Conference on Acoustics Speech and Signal Processing, Toulouse, France, 14–19 May 2006; Volume 3. [CrossRef]
23. Isvoranu, D.; Udrişte, C. Fluid flow versus geometric dynamics. In Proceedings of the 5th Conference on Differential Geometry, Mangalia, Romania, 29 August–2 September 2005; Geometry Balkan Press: Bucharest, Romania, 2006; pp. 70–82.
24. van der Pol, B. On "relaxation-oscillations". *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1926**, *2*, 978–992. [CrossRef]
25. Tôrres, L.; Aguirre, L. Transmitting information by controlling nonlinear oscillators. *Physica D* **2004**, *196*, 387–406. [CrossRef]
26. Cao, T.; Yi, H. On the complex oscillation of higher order linear differential equations with meromorphic coefficients. *J. Syst. Sci. Complex.* **2007**, *20*, 135–148. [CrossRef]
27. Chen, J.; Lu, J.; Wu, X. Bidirectionally coupled synchronization of the generalized Lorenz systems. *J. Syst. Sci. Complex.* **2011**, *24*, 433–448. [CrossRef]
28. Huang, J.; Zhang, H. Bifurcations of periodic orbits in three-well Duffing system with a phase shift. *J. Syst. Sci. Complex.* **2011**, *24*, 519–531. [CrossRef]
29. Mo, J.; Lin, W. Generalized variation iteration solution of an atmosphere-ocean oscillator model for global climate. *J. Syst. Sci. Complex.* **2011**, *24*, 271–276. [CrossRef]
30. Fiori, S. Nonlinear damped oscillators on Riemannian manifolds: Fundamentals. *J. Syst. Sci. Complex.* **2016**, *29*, 22–40. [CrossRef]
31. FitzHugh, R. Mathematical models of excitation and propagation in nerve. In *Biological Engineering*; H.P. Schwan, Ed.; McGraw-Hill: New York, NY, USA, 1969.
32. Barbosa, R.; Tenreiro Machado, J.; Vinagre, B.; Calderón, A. Analysis of the van der Pol oscillator containing derivatives of fractional order. *J. Vib. Control* **2007**, *13*, 1291–1301. [CrossRef]

33. Duffing, G. *Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz und Ihre Technische Bedeutung*; Braunschweig, F., Ed.; Vieweg & Sohn: Berlin, Germany, 1918.
34. Trueba, J.; Rams, J.; Sanjuán, M. Analytical estimates of the effect of nonlinear damping in some nonlinear oscillators. *Int. J. Bifurc. Chaos* **2000**, *10*, 2257–2267. [CrossRef]
35. Sanjuán, M. The effect of nonlinear damping on the universal escape oscillator. *Int. J. Bifurc. Chaos* **1999**, *9*, 735–744. [CrossRef]
36. Lorenz, E. Deterministic nonperiodic flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
37. Greer, M.; Saha, R.; Gogliettino, A.; Yu, C.; Zollo-Venecek, K. Emergence of oscillations in a simple epidemic model with demographic data. *R. Soc. Open Sci.* **2020**, *7*, 191187. [CrossRef]
38. Georgiou, I.; Corless, M.; Bajaj, A. Dynamics of nonlinear structures with multiple equilibria: A singular perturbation-invariant manifold approach. *Z. Angew. Math. Phys.* **1999**, *50*, 892–924. [CrossRef]
39. Shorek, S. A stationarity principle for non-conservative systems. *Adv. Water Resour.* **1984**, *7*, 85–88. [CrossRef]
40. Molero, F.; Lara, M.; Ferrer, S.; Cèspedes, F. 2-D Duffing oscillator: Elliptic functions from a dynamical systems point of view. *Qual. Theory Dyn. Syst.* **2013**, *12*, 115–139. [CrossRef]
41. Cucker, F.; Smale, S. Emergent behavior in flocks. *IEEE Trans. Autom. Control* **2007**, *52*, 852–862. [CrossRef]
42. Ha, S.Y.; Kim, D.; Schlöder, F.W. Emergent behaviors of Cucker–Smale flocks on Riemannian manifolds. *IEEE Trans. Autom. Control* **2021**, *66*, 3020–3035. [CrossRef]
43. Fiori, S.; Del Rossi, L. Minimal control effort and time Lie-group synchronisation design based on proportional-derivative control. *Int. J. Control* **2022**, *95*, 138–150. [CrossRef]
44. Fiori, S. Synchronization of first-order autonomous oscillators on Riemannian manifolds. *Discret. Contin. Dyn. Syst. Ser. B* **2019**, *24*, 1725–1741. [CrossRef]
45. Fiori, S. Model formulation over Lie groups and numerical methods to simulate the motion of gyrostats and quadrotors. *Mathematics* **2019**, *7*, 935. [CrossRef]
46. Gaponov, I.; Razinkova, A. Quadcopter design and implementation as a multidisciplinary engineering course. In Proceedings of IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE) 2012, Hong Kong, China, 20–23 August 2012; pp. H2B-16–H2B-19. [CrossRef]
47. Bloch, A.; Krishnaprasad, P.; Marsden, J.; Ratiu, T. The Euler-Poincaré equations and double bracket dissipation. *Commun. Math. Phys.* **1996**, *175*, 1–42. [CrossRef]
48. Gajbhiye, S.; Banavar, R. The Euler-Poincaré equations for a spherical robot actuated by a pendulum. *IFAC Proc. Vol.* **2012**, *45*, 72–77. [CrossRef]
49. Becker, M.; Sampaio, R.; Bouabdallah, S.; de Perrot, V.; Siegwart, R. In-flight collision avoidance controller based only on OS4 embedded sensors. *J. Braz. Soc. Mech. Sci. Eng.* **2012**, *34*, 295–297. [CrossRef]
50. Fiori, S. A control-theoretic approach to the synchronization of second-order continuous-time dynamical systems on real connected Riemannian manifolds. *SIAM J. Control Optim.* **2020**, *58*, 787–813. [CrossRef]
51. Magdy, M.; Ng, T. Regulation and control effort in self-tuning controllers. *IEE Proc. D Control Theory Appl.* **1986**, *133*, 289–292. [CrossRef]
52. Grönwall, T. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. Math.* **1919**, *20*, 292–296. [CrossRef]
53. Henderson, H.; Searle, S. On deriving the inverse of a sum of matrices. *SIAM Rev.* **1981**, *23*, 53–60. [CrossRef]
54. Gavrilov, A. The Taylor series related to the differential of the exponential map. *arXiv* **2012**, arXiv:1205:2868v1.
55. Bullo, F.; Lewis, A. *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Mechanical Control Systems*; Springer: Berlin/Heidelberg, Germany, 2005.
56. Osborne, J.; Hicks, G. The geodesic spring on the Euclidean sphere with parallel-transport-based damping. *Not. AMS* **2013**, *60*, 544–556. [CrossRef]
57. Ambrose, W.; Singer, I.M. A theorem on holonomy. *Trans. Am. Math. Soc.* **1953**, *75*, 428–443. [CrossRef]
58. Fiori, S. Auto-regressive moving-average discrete-time dynamical systems and autocorrelation functions on real-valued Riemannian matrix manifolds. *Discret. Contin. Dyn. Syst. Ser. B* **2014**, *19*, 2785–2808. [CrossRef]
59. Escobar, J.; Poznyak, A. Robust parametric identification for ARMAX models with non-Gaussian and coloured noise: A survey. *Mathematics* **2022**, *10*, 1291. [CrossRef]