**ERROR-BASED CONTROL SYSTEMS ON RIEMANNIAN STATE MANIFOLDS: PROPERTIES OF THE PRINCIPAL PUSHFORWARD MAP ASSOCIATED TO PARALLEL TRANSPORT**

**SIMONE FIORI**  
Universit`a Politecnica delle Marche  
Via Brecce Bianche, I-60131 Ancona, Italy

(Communicated by Yacine Chitour)

**Abstract.** The objective of the paper is to contribute to the theory of error-based control systems on Riemannian manifolds. The present study focuses on systems where the control field influences the covariant derivative of a control path. In order to define error terms in such systems, it is necessary to compare tangent vectors at different points using parallel transport and to understand how the covariant derivative of a vector field along a path changes after such field gets parallely transported to a different curve. It turns out that such analysis relies on a specific map, termed principal pushforward map. The present paper aims at contributing to the algebraic theory of the principal pushforward map and of its relationship with the curvature endomorphism of a state manifold.

1. **Introduction.** The objective of this paper is to contribute to the theory of error-based control systems on Riemannian manifolds. A frequently-occurring example is given by the regulation of the dynamics of unmanned aerial or extra-terrestrial vehicles (such as drones or satellites), whose attitude may be described by a three-dimensional rotation matrix with respect to a earth-fixed reference frame: their dynamics may be described in terms of second-order state-space-type non-linear differential equations on the special orthogonal group SO(3) [16, 31], which is a three-dimensional smooth manifold.

More specifically, in this paper we focus on systems where the control \( u \) influences the covariant derivative \( \nabla_x \dot{x} \) of a control path \( x(t) = x_u(t) \). In order to define error terms in such systems, one will need to compare tangent vectors at different points using parallel transport along minimal geodesics \( \mathcal{P}^{y \to x} : T_y M \to T_x M \), and furthermore, to understand how the covariant derivative of a vector field \( v_x(t) := \mathcal{P}^{y(t) \to x(t)}(w_y(t)) \) depends on the covariant derivative of \( w_y(t) \) as well as derivatives of the curves \( y(t) \) and \( x(t) \) involved. This paper shows how such covariant derivatives are related through what is called the principal pushforward map \( d_x \mathcal{P}^{y \to x} \). We show that this map is the only ingredient needed to relate the covariant derivatives of \( w_y(t) \) and \( v_x(t) \). Furthermore, we shown that this map may be written as an

---

2020 *Mathematics Subject Classification.* Primary: 93C10, 93C25, 58C25.

*Key words and phrases.* Error-based non-linear dynamical systems, Riemannian state manifolds, Parallel transport, Covariant differentiation, Principal pushforward map, Curvature endomorphism.
integral depending on the curvature of the manifold. The results conveyed in this paper possess interesting applications to non-linear control, such as in the solution of the problem of synchronization of Duffing oscillators.

The paper is written in general terms and does not assume a deep knowledge of differential geometry. To keep the paper suitable for a general audience, the main concepts (both those surveyed from the specific literature and the main findings) are either proven (whenever necessary) and illustrated by examples. In fact, all the sections of the paper are enriched by a number of examples, related to the unit hypersphere, that help clarifying their theoretical content and provide readers a glimpse about the practical involved calculations.

This paper is organized as follows. Section 2 recalls some notions from manifold calculus, such as the definition of the parallel transport operator and of the Riemannian covariant derivative. This section also introduces a methodological discussion to motivate the present study. Section 3 defines and describes some algebraic properties of the principal pushforward map. Section 4 discusses in detail a case of study related to the synchronization of two non-linear Duffing-type oscillators on Riemannian manifolds. Section 5 elucidates a relationship between the principal pushforward map and the curvature endomorphism.

2. Parallel transport, covariant derivative, motivations. In the present paper, it is assumed that $M$ is a real geodesically complete Riemannian manifold. The symbol $TM$ stands for the tangent bundle associated to the manifold $M$, which is supposed to be endowed with the Riemannian connection $\nabla$. For dynamical systems, a manifold plays the role of state space, while the associated tangent bundle plays the role of phase-space [2]. The present section is divided in two parts: The Subsection 2.1 recalls the formal definition of parallel transport operator, along with some of its prominent features, and its relationship with covariant differentiation on a Riemannian manifold; this subsection also establishes the notation used within the paper. The Subsection 2.2 motivates the content of the present paper by mentioning a case of study, namely a tracking problem for paired dynamical systems, where the algebraic properties of the parallel transport operator are instrumental in studying tracking ability.

2.1. Parallel transport operator and Riemannian covariant derivative. The notion of parallel transport plays a central role in a variety of theoretical [21, 28] as well as applied problems [9, 25] and, in particular, provides a way of studying and simulating numerically the behavior of dynamical systems on smooth curved manifolds [12, 14, 24]. Parallel transport is a way of transporting geometrical objects along piecewise smooth state-trajectories in a smooth state manifold $M$ [5], as it supplies a way of moving the local geometry of a manifold along a curve, namely, of connecting the geometries of nearby points [29]. As parallel transport supplies a local realization of the connection, it also supplies a local realization of the curvature known as holonomy.

Let $\gamma : I \rightarrow M$ be a smooth curve from an interval $I$ into $M$ and let $v_0 \in T_{\gamma(t_0)}M$ be a vector tangent to $M$ at $\gamma(t_0)$ for some $t_0 \in I$. A tangent vector field $v_\gamma(t)$ is generated by parallel transport of $v_0$ along the curve $\gamma(t)$ provided that $v_\gamma(t)$, for all $t \in I$, is a tangent vector for which $v_\gamma(t_0) = v_0$. The qualifier “parallel” refers to the fact that the vector field $v_\gamma(t)$ along a curve is necessarily covariantly constant, i.e., $v_\gamma(t)$ satisfies the equation $\nabla_{\dot{\gamma}(t)} v_\gamma = 0$ for all $t \in I$. A standard result in differential geometry is that, under smoothness hypotheses, parallel transport is unique [29].
A distinguishing feature of parallel transport is that it defines a linear isomorphism between tangent spaces. In fact, fix two values \( t_0, t \in \mathbb{I} \) and consider the parallel transport of the tangent vector \( v_0 \in T_{t_0(t_0)}\mathbb{M} \) to \( v_\gamma(t) \in T_{\gamma(t)}\mathbb{M} \) along the curve \( \gamma \). The tangent vector \( v_\gamma(t) \) is then obtained by a linear transformation \( \tau_\gamma^t : T_{t_0(t_0)}\mathbb{M} \to T_{\gamma(t)}\mathbb{M} \), namely, \( v_\gamma(t) = \tau_\gamma^t(v_0) \). The transformation \( \tau_\gamma^t \) is clearly invertible, its inverse \( (\tau_\gamma^t)^{-1} \) being given by the parallel transport along the reversed arc of \( \gamma \) from \( t \) to \( t_0 \), and the linear transformation \( \tau_\gamma^0 \) coincides with the identity \( \text{id}_{\gamma(t_0)} \), where \( \text{id}_x \) stands for \( \text{id}_{T_x\mathbb{M}} \).

In the present contribution, we are interested in a special instance of parallel transport. In fact, we assume that the finite-dimensional Riemannian manifold \( \mathbb{M} \) is geodesically complete, namely, that geodesics can be extended indefinitely. Geodesically complete manifolds are also geodesically convex, which means that any two points \( x, y \in \mathbb{M} \) may be connected by a geodesic arc \( \gamma_{xy} : [0, 1] \to \mathbb{M} \) such that \( \gamma_{xy}(0) = x \) and \( \gamma_{xy}(1) = y \) (the right arrow over the endpoints signifies that the geodesic arc is traversed from \( x \) to \( y \)). Since, under smoothness conditions and provided that the endpoints are sufficiently close to one another, the geodesic connecting such two endpoints is unique, the corresponding parallel transport \( \tau_{\gamma_{xy}} \) is uniquely defined by the endpoints. Therefore, the operator \( \mathcal{P}^{x \to y} : T_x\mathbb{M} \to T_y\mathbb{M} \) is defined to represent the parallel transport of a tangent vector at a point \( x \in \mathbb{M} \) to the tangent space at a point \( y \in \mathbb{M} \) along the geodesic arc connecting \( x \) to \( y \), namely \( \mathcal{P}^{x \to y} := \tau_{\gamma_{xy}} \). The notion of parallel transport is illustrated in the following example related to the unit hypersphere.

**Example 1.** Consider the special case that the manifold \( \mathbb{M} \) is the unit-radius hypersphere \( S^{n-1} \) embedded in the ambient space \( A := \mathbb{R}^n \). The unit hypersphere and its tangent spaces read:

\[
S^{n-1} := \{ x \in A \mid x^T x = 1 \},
\]

\[
T_x S^{n-1} = \{ v \in A \mid x^T v = 0 \}.
\]

Assume that the manifold \( S^{n-1} \) is endowed with the metric \( \langle v, w \rangle_x := v^T w \) inherited from the Euclidean metric in the ambient space \( A \), for any \( v, w \in T_x S^{n-1} \). The parallel transport operator is expressed by (adapted from [11]):

\[
\mathcal{P}^{x \to y}(v) = \left[ I_n - \frac{(y + x)(y^T)}{1 + x^T y} \right] v,
\]

where it is assumed that the points \( x, y \in S^{n-1} \) are not antipodal. The symbol \( I_n \) denotes a \( n \times n \) identity matrix.

Two properties of the parallel-transport operator that may be easily proven directly are that \( \mathcal{P}^{x \to x} \) is an identity and that the vector \( \mathcal{P}^{x \to y}(v) \) is tangent at \( y \):

\[
\mathcal{P}^{x \to x}(v) = \left[ I_n - xx^T \right] v = v,
\]

\[
y^T \mathcal{P}^{x \to y}(v) = \left[ y^T - (1 + x^T y)^{-1}(1 + y^T x)y^T \right] v = (y - y)^T v = 0,
\]

for every \( x \in S^{n-1}, y \in S^{n-1} \setminus \{-x\} \) and \( v \in T_x S^{n-1} \). Indeed, the linear operator \( \mathcal{P} \) may be extended to the whole ambient space \( A \), in which it admits the matrix representation

\[
\mathbb{P}^{x \to y} := I_n - \frac{(y + x)y^T}{1 + x^T y},
\]

in the sense that \( \mathcal{P}^{x \to y}(v) = \mathbb{P}^{x \to y}v \). From here, it is easy to see that the matrix \( \mathbb{P}^{x \to y} \) represents an orthogonal projection onto \( T_y S^{n-1} \), in that \( \mathbb{P}^{x \to y}^2 = \mathbb{P}^{x \to y} \).
Further examples related specifically to the ordinary sphere $S^2$ can be found in [24]. Since the connection $\nabla$ that the parallel transport $\mathcal{P}$ is defined from is Riemannian, the parallel transport realizes an isometry and is referred to as metric transport. Namely, given two points $x, y \in M$ and two tangent vectors $u, v \in T_xM$, it holds that

$$\langle \mathcal{P}_{x \rightarrow y}(u), \mathcal{P}_{x \rightarrow y}(v) \rangle_y = \langle u, v \rangle_x. \quad (5)$$

In fact, a metric parallel transport does not change the angle nor the length of the transported vectors, and this is equivalent to the connection being compatible with the metric. The isometry property of parallel transport is illustrated in the following example related to the unit hypersphere.

**Example 2.** For the case of the hypersphere endowed with the Euclidean inner product, the isometry property is easy to verify directly:

$$[\mathcal{P}_{x \rightarrow y}(u)]^T \mathcal{P}_{x \rightarrow y}(v) = u^T \left[ I_n - \frac{y(y + x)^T}{1 + x^T y} \right] \left[ I_n - \frac{y(y + x)^T}{1 + x^T y} \right] v
= u^T v - \frac{u^T (x + y)(y^T v)}{1 + x^T y} - \frac{(u^T y)(x + y)^T v}{1 + x^T y}
+ \frac{(u^T y)(x + y)^T (x + y)(y^T v)}{(1 + x^T y)^2}
= u^T v - \frac{(u^T y)(v^T y)}{1 + x^T y} - \frac{(u^T y)(v^T y)}{1 + x^T y}
+ \frac{(u^T y)(2 + 2x^T y)(v^T y)}{(1 + x^T y)^2}
= u^T v,$$

which is an instance of the general property (5). In particular, the above identity shows that $\|\mathcal{P}_{x \rightarrow y}(u)\| = \|u\|$.

Upon defining the angle between two vectors $u, v \in T_xM$ as

$$\alpha(u, v) := \cos \left( \frac{u^T v}{\|u\| \|v\|} \right),$$

where $\|u\| := \sqrt{\langle u, u \rangle}$ denotes the Fröbenius norm induced by the inner product and ‘acos’ denotes the inverse cosine function, from the isometry property it is immediate to verify that

$$\alpha(P_{x \rightarrow y}(u), P_{x \rightarrow y}(v)) = \acos \left( \frac{[P_{x \rightarrow y}(u)]^T P_{x \rightarrow y}(v)}{\|P_{x \rightarrow y}(u)\| \|P_{x \rightarrow y}(v)\|} \right)
= \acos \left( \frac{u^T v}{\sqrt{u^T u} \sqrt{v^T v}} \right)
= \alpha(u, v),$$

namely, the angle between the transported vectors is preserved.

From now on, we will denote by $x(t)$ a smooth curve in $M$ that originates at a point $x$ and we let $\dot{x}(t)$ be its velocity field, with $t$ belonging to an interval $I \subset \mathbb{R}$. (To be consistent with the previous notation, we should write $\dot{x}_x(t)$, however, the subscript $x$ would be pleonastic). The covariant derivative of a tangent vector field
denotes the naive directional derivative of a multi-variable function in $\mathbb{R}^n$, properly corrected for the fact that $v_x(t)$ and $v_x(t + s)$ belong to different tangent spaces and need therefore to be aligned – by means of the parallel transport – before being compared to one another.

**Example 3.** For the case of the hypersphere endowed with the Euclidean inner product, the covariant derivative associated to the parallel transport operator recalled in (3) may be obtained as follows:

$$
\lim_{s \to 0} \frac{\partial x(t + s) \to x(t)}{s} (v_x(t + s)) - v_x(t)
$$

which represents the rate of change of the vector field $v_x$ in the direction of the velocity field $\dot{x}$. The above formula clearly represents the analogous of the directional derivative of a multi-variable function in $\mathbb{R}^n$, properly corrected for the fact that $v_x(t)$ and $v_x(t + s)$ belong to different tangent spaces and need therefore to be aligned – by means of the parallel transport – before being compared to one another.

$$
\llbracket I_n - \frac{x(t + s) + x(t)}{1 + x^T(t)x(t + s)} \rrbracket v_x(t + s) - v_x(t) = 0
$$

where we made use of the property $x^T(t + s) v_x(t + s) = 0$ and where

$$
v_x'(t) := \lim_{s \to 0} \frac{v_x(t + s) - v_x(t)}{s} \in \mathbb{A}
$$

denotes the naive directional derivative of the vector field $v$ along $\dot{x}$ in the ambient space $\mathbb{A}$ (namely, in general, $v_x'(t) \notin T_x S^{n-1}$). The above equations chain shows that

$$
\nabla_\dot{x} v_x = v_x' + x \dot{x}^T v_x.
$$

If one takes $v_x = \dot{x}$, then $\nabla_\dot{x} \dot{x} = \dot{x}' + x \dot{x} = \ddot{x},$ which is used to calculate the expression of geodesic arcs on the unit hypersphere.

It is interesting to verify directly that the covariant derivative (8) returns a tangent vector at $x$. To achieve a proof, let us recall that $x^T(t) v_x(t) = 0$ for every value of $t \in I$. Deriving both sides of the preceding equation, we get $\dot{x}^T v_x = -x^T v_x'$, which gives the alternative expression for the covariant derivative

$$
\nabla_\dot{x} v_x = v_x' + x (-x^T v_x') = (I_n - x x^T) v_x'.
$$

The last expression confirms the intuitive meaning of covariant derivative as orthogonal projection of the naive derivative onto the tangent space $T_x S^{n-1}$ and confirms, at the same time, that the covariant derivative returns a vector in $T_x S^{n-1}$.

**2.2. Motivation of the present research.** A large corpus of linear and non-linear system dynamics theory is based on the evaluation of an error figure, as in tracking systems, where the error measures the discrepancy between the actual system state and its predicted value $[18, 33]$, in leader/follower system synchronization, where the error measures the deviation between the state of the leader and the state of the follower $[19, 20]$, in system state observation, where the error quantifies the
disagreement between the actual state of the observed system and the estimated state by an observer \([10, 17, 30]\), and in system control/regulation, where the error measures the incongruity between the actual state of a system and its desired value (or set point) and provides a feedback information to the controller.

Given two dynamical systems whose state spaces are linear subspaces of the Euclidean space \(\mathbb{R}^n\) and whose states are represented by the vector-valued variables \(x, y \in \mathbb{R}^n\), a vector-valued error is typically defined as \(e(t) := x(t) - y(t)\). By means of vector calculus, it is then easy to compute the time-derivatives \(\dot{e} := \frac{dx}{dt} = \dot{x} - \dot{y}\) and \(\ddot{e} := \frac{d^2e}{dt^2} = \ddot{x} - \ddot{y}\) that may be used to quantify the time-evolution of the error and to study its asymptotic dynamics, by means of what is termed error system. This is a specific terminology that indicates a differential equation that the error function \(e(t)\) must satisfy and whose qualitative analysis provides valuable information on the asymptotic behavior of the systems in relation to one another.

**Example 4.** As a simple example, let us consider the mass-spring-damper system depicted in the Figure 1, where it is assume that, in a fixed mono-dimensional coordinate system, the position of the mass \(m\) is denoted by \(x(t)\) and the position of the reference point is denoted by \(y(t)\); notice that in a tracking problem, the reference point may be moving over time. Let us denote by \(k_D > 0\) the stiffness of the damper and by \(k_P\) the elasticity constant of the spring. The Newton’s law of dynamics for the mass \(m\) writes, then:

\[
m \frac{d^2(x - y)}{dt^2} + k_D \frac{d(x - y)}{dt} + k_P (x - y) = 0.
\]

(10)

Since the difference \(x(t) - y(t)\) corresponds to the tracking error \(e(t)\), the above equation gives directly rise to the tracking-error-system

\[
m\ddot{e} + k_D \dot{e} + k_P e = 0,
\]

(11)

which is a second-order differential equation in the time-function \(e(t)\) that may be studied in details, for example by an energy-dissipation analysis, to prove that the tracking error tends asymptotically to zero.

Speaking, instead, of two dynamical systems whose common state space is a curved Riemannian state manifold \(\mathbb{M}\), whose states are described by the manifold-valued variables \(x, y \in \mathbb{M}\), there is not a straightforward definition of error, although the paper \([24]\), based on the previous study \([4]\), defines the error term \(e\) as the Riemannian gradient of the squared Riemannian distance between the two states. Here, we shall assume that the Riemannian gradient is taken with respect to the state \(x\), in such a way that \(e \in T_x \mathbb{M}\), namely, the error term is a tangent vector field along the trajectory of one of the systems. What is most relevant to the present
discussion, is that, as a result of such definition, the error evolution is described by
the adjoint error term \( \varepsilon := \dot{x} - \mathcal{P}y(t) \rightarrow x(t)(\dot{y}) \) and by the covariant derivative \( \nabla_x \varepsilon \),
which affords defining an error dynamics for the error terms.

Example 5. The mass-spring-damper-based system recalled in the previous example
was extended to a tracking problem on a manifold (the unit hypersphere) in [24],
where it was shown that the error term satisfies the error system

\[
\nabla \varepsilon + \kappa_D \varepsilon + \kappa_P e = 0
\]

which, by an appropriate Lyapunov-stability analysis [24], affords proving that
the tracking system is asymptotically stable. As it is very apparent, the er-
ror system (12) is based on the evaluation of the covariant derivative

\[
\nabla \varepsilon = \nabla \dot{x} - \nabla_y \mathcal{P}y(t) \rightarrow x(t)(\dot{y}).
\]

While the first term on the right-hand side is nothing but the \textit{geometric acceleration}
of the state \( x \), the second term is more involved and
needs a detailed investigation.

A more involved problem, namely the synchronization of two non-linear Duffing-
type oscillators on Riemannian manifolds, will be discussed in details by applying
the main results of the paper.

The present paper is devoted to the analysis of the quantity \( \nabla \mathcal{P}y(t) \rightarrow x(t)(\dot{y}) \).

The analysis carried out in the following sections elucidates some interesting alge-
braic properties of the principal pushforward map associated to the parallel trans-
port operator and the connection between such principal pushforward map and the
curvature endomorphism of the state manifold.

3. The \textbf{principal pushforward map of parallel transport and its main}
properties. Consider the following problem which, as illustrated in Subsection 2.2,
arises from the the motivation to the present study. Given two smooth trajectories
on a connected Riemannian state manifold and a tangent vector field attached to
one of the trajectories like, for example, the velocity vector field, it may occur to
be necessary to parallel-transport this vector field to the other trajectory to get
a secondary vector field, as schematically depicted in Figure 2. A question arises
naturally, namely \textit{how do the rates of change of these two vector fields relate to one
another?}

The answer to this question involves two maps, that are referred to as \textit{partial
pushforward maps}, which are defined – and commented – in Subsection 3.1. Since
one of them may be expressed in terms of the other, the latter will be referred to as
\textit{principal pushforward map}. The main properties of the principal pushforward map
are analyzed in Subsection 3.2.

3.1. \textbf{Definition of principal pushforward map.} The \textit{partial pushforward maps}
associated to the parallel transport operator are defined as follows:

\textbf{Definition 3.1.} Let \( f : \mathbb{M} \times T\mathbb{M} \to T\mathbb{M} \) be defined as \( f(x, y, w_y) := \mathcal{P}y \rightarrow x(w_y) \)
for any two smooth curves \( x(t) \) and \( y(t) \) in \( \mathbb{M} \) and for any tangent vector field
\( w_y(t) \in T_{y(t)}\mathbb{M} \). The partial pushforward maps of the function \( f \) are defined as

\[
\begin{align*}
(d_x f)_{x(t), y(t), w_y(t)}(\dot{x}) & := \lim_{s \to 0} \frac{1}{2} [\mathcal{P}x(t+s) \rightarrow x(t)(f(x(t+s), y(t), w_y(t))) \\
& - f(x(t), y(t), w_y(t))], \\

(d_y f)_{x(t), y(t), w_y(t)}(\dot{y}) & := \lim_{s \to 0} \frac{1}{2} [f(x(t), y(t+s), \mathcal{P}y(t) \rightarrow y(t+s)(w_y(t))) \\
& - f(x(t), y(t), w_y(t))],
\end{align*}
\]

(13)
where \(d_x f\) denotes the partial pushforward map of the function \(f\) with respect to its first argument, while the symbol \(d_y f\) denotes the partial pushforward map of the function \(f\) with respect to its second argument.

The notation \((d_x f)_{x,y,w}\) means that this pushforward map, which is a function of three values (not to be confused with its argument, that is a fourth, tangent-space-valued variable), is evaluated at a point where the first argument is instantiated to the value \(x\), the second argument is instantiated to the value \(y\) and the third argument is instantiated to the value \(w\).

Since the map \((d_y f)_{x,y,w}\) may be expressed in terms of the map \((d_x f)_{x,y,w}\), the latter will be referred to as principal pushforward map. This map acts more like a kind of covariant derivative than an actual differential, as it will be shown in the next subsection and, in particular, in Example 6.

### 3.2. Main properties of the principal pushforward map.

The principal pushforward map of the parallel transport are related to the covariant derivative of a parallely-transported vector field by the following result.

**Theorem 3.2.** Take two smooth curves \(x(t)\) and \(y(t)\) in \(\mathbb{M}\) and a smooth vector field \(w_y(t) \in T_y M\) and define the new vector field \(v_x(t) \in T_x M\) by

\[
v_x(t) := P_{y \to x}(w_y(t)), \tag{14}
\]

with \(t \in \mathbb{I}\). The covariant derivative of \(v_x\) along the velocity field of the curve \(x(t)\) is computed as

\[
\nabla_x v_x = (d_x \mathcal{P}^{y \to x})_{x,y,w_y}(\dot{x}) - \mathcal{P}^{y \to x}[ (d_x \mathcal{P}^{y \to x})_{y,x,v_x} ](\dot{y}) - \nabla_y w_y, \tag{15}
\]

where \((d_x \mathcal{P}^{y \to x})_{x,.,.} \in \text{End}(T_x \mathbb{M})\) for any \(x \in \mathbb{M}\) is termed principal pushforward map.

**Proof.** Let \(f : \mathbb{M} \times T \mathbb{M} \to T \mathbb{M}\) be defined as \(f(x, y, w_y) := \mathcal{P}^{y \to x}(w_y)\) for any two points \(x\) and \(y\) in \(\mathbb{M}\) and for any tangent vector \(w_y \in T_y \mathbb{M}\). Note, in particular, that the function \(f\) is linear in the third argument. Now, set \(v_x(t) := f(x(t), y(t), w_y(t))\)
and note that \(v_x(t) \in T_x(M)\). By definition of covariant derivative along a curve, it holds that
\[
\nabla_{\dot{x}}v_x = \lim_{s \to 0} \frac{1}{s} \left[ f_{x(t+s)}(v_x(t+s)) - v_x(t) \right]
\]
where \(\dot{x}\) is the principal pushforward map. By definition of covariant derivative along a curve, the identity (17), the result (15) is obtained.

Now, since \(\mathcal{P}^{x\to y} \circ \mathcal{P}^{y\to z} = \text{id}_z\), the following identity holds:
\[
w_y = f(y, x, f(x, y, w_y)).
\]

Deriving both sides with respect to the covariant derivative \(\nabla\), we get the identity
\[
\nabla_yw_y = (d_xf)_{y,x,v_y}(\dot{y}) + (d_yf)_{y,x,v_y}(\dot{x}) + f(y, x, (d_xf)_{x,y,w_y}(\dot{x})
\]
+ \((d_yf)_{x,y,w_y}(\dot{y}) + f(x, y, \nabla_yw_y))\).

Notice that \((d_xf)_{y,x,v_x} \in \text{End}(T_yM)\), while \((d_yf)_{x,y,w_y}\) maps \(T_yM\) to \(T_xM\). The above identity is equivalent to
\[
\nabla_yw_y = (d_xf)_{y,x,v_x}(\dot{y}) + (d_yf)_{y,x,v_x}(\dot{x}) + f(y, x, (d_xf)_{x,y,w_y}(\dot{x})
\]
+ \((d_yf)_{x,y,w_y}(\dot{y}) + \nabla_yw_y)\).

Since the velocity vector fields \(\dot{x}\) and \(\dot{y}\) are arbitrary, the above formula yields two independent operator identities, one of which is \(f(y, x, (d_yf)_{y,x,w_y} + (d_xf)_{y,x,f(x,y,w_y}) = 0\), namely
\[
(d_yf)_{x,y,w_y} = -f(x, y, (d_xf)_{x,y,f(x,y,w_y)}).
\]

Since the pushforward map \(d_yf\) may be written in terms of the pushforward map \(d_xf\), the latter is termed principal pushforward map. By using the equation (16), the definition of \(f\) and the identity (17), the result (15) is obtained.

A couple of examples would serve to clarify the motivation behind the above analysis and how the partial pushforward maps may be calculated on a given manifold.

**Example 6.** Assume that one of the two curves \(x(t)\) and \(y(t)\) considered in the statement of Theorem 3.2 is constant, for example, suppose that
\[
y(t) = y = \text{constant for every } t \in \mathbb{I},
\]
in such a way that its image is a single point \(\{y\}\). As a consequence, we have that \(\dot{y} = 0\) and therefore \(\nabla_{\dot{y}}w_y = 0\) and \((d_x\mathcal{P}^{y\to x})_{x,y,v_x}(\dot{y}) = 0\). Nevertheless, the vector
field \( v_x(t) = \mathcal{P}^{y \to x(t)}(w_y) \) changes over time and its covariant derivative along the velocity field \( \dot{x} \) is still given by the relation (15) that, by the hypothesis made, reads:
\[
(d_x \mathcal{P}^{y \to x})_{x,y,w_y}(\dot{x}) = \nabla_{\dot{x}} \mathcal{P}^{y \to x}(w_y) \big|_{y(t) = \text{const}}. \tag{19}
\]
The above relation provides a meaning of the principal pushforward map \((d_x \mathcal{P}^{y \to x})_{x,y,w_y}\) as the covariant derivative of the transported constant vector field \(w_y\).

Example 7. Consider the parallel translation map for the unit hypersphere corresponding to a Euclidean metric. By applying the definition of pushforward maps in terms of time-derivatives, it is straightforward to compute the maps \((d_x \mathcal{P}^{y \to x})_{x,y,w_y}\) and \((d_y \mathcal{P}^{y \to x})_{x,y,w_y}\).

Let us start by recalling that, in the present example,
\[
f(x, y, w_y) := \left[ I_n - \frac{(y + x)x^T}{1 + x^Ty} \right] w_y
\]
The naïve derivative of this function reads
\[
f' = \left[ -\frac{(\dot{y} + \dot{x})x^T + (y + x)\dot{x}^T(1 + x^Ty) - (y + x)x^T(x^T\dot{y} + x^T\dot{y})}{(1 + x^Ty)^2} \right] w_y
\]
\[+ f(x, y, w'_y).\tag{20}\]

On the other hand, by using the expression of the covariant derivative of the vector fields \(v_x = f(x, y, w_y)\) and \(v_y = f(x, y, w'_y)\) given in (8), the relation (16) can be rewritten as:
\[
f' + x f^T \dot{x} = (d_x f)_{x,y,w_y}(\dot{x}) + (d_y f)_{x,y,w_y}(\dot{y}) + f(x, y, w'_y + yw^T \dot{y}). \tag{21}\]

Plugging the equation (20) into the relation (21) gives
\[
-\frac{(\dot{y} + \dot{x})(x^Tw_y) + (y + x)(x^Tw_y) + (y + x)(x^Tw_y)(\dot{x}^Ty + x^T\dot{y})}{(1 + x^Ty)^2}
\]
\[+ xw^T_y \left[ I_n - \frac{(y + x)x^T}{1 + x^Ty} \right] \dot{x} = (d_x f)_{x,y,w_y}(\dot{x}) + (d_y f)_{x,y,w_y}(\dot{y})
\]
\[+ (w^T \dot{y}) \left[ I_n - \frac{(y + x)x^T}{1 + x^Ty} \right] y.\]

By separating the terms in \(\dot{x}\) and \(\dot{y}\), after some lengthy calculations we get a matrix representation of the sought pushforward maps
\[
[(d_y \mathcal{P}^{y \to x})_{x,y,w_y}] = -\frac{x^T w_y (I_n + x y^T)}{1 + x^T y} + x w^T_y
\]
\[+ \frac{(x^T w_y) (y + x) y^T}{(1 + x^T y)^2}, \tag{22}\]
\[
[(d_y \mathcal{P}^{y \to x})_{x,y,w_y}] = -\frac{x^T w_y}{1 + x^T y} \left[ I_n - \frac{(y + x) x^T}{1 + x^T y} \right] - y w^T_y
\]
\[+ \frac{(x^T y) (y + x) w^T_y}{1 + x^T y}. \tag{23}\]

Both pushforward functions map onto \(T_x M\), namely
\[
x^T (d_x \mathcal{P}^{y \to x})_{x,y,w_y}(\dot{x}) = x^T (d_y \mathcal{P}^{y \to x})_{x,y,w_y}(\dot{y}) = 0.
\]
Moreover, calculations allow one to verify the identity (17), that may be rewritten explicitly as \((d_x \mathcal{P}^{y \rightarrow x})_{x,y,w_y}(\dot{y}) = -\mathcal{P}^{y \rightarrow x} \circ (d_x \mathcal{P}^{y \rightarrow x})_{y,x,\mathcal{P}^{x \rightarrow y}(w_x)}(\dot{y})\), where the symbol \(\circ\) denotes the usual composition of maps.

It is instructive to derive an expression of the principal pushforward map in local coordinates\(^1\), as outlined in the following

**Corollary 1.** Chosen a local coordinate system, denote by \(\{\partial_i\}_x\) the basis of \(T_x \mathcal{M}\) with respect to the local coordinates, by \(\{d^i\}_x\) the basis of the dual space \(T^*_x \mathcal{M}\), by \(\Gamma_{ij}^k\) the Christoffel symbols of the Riemannian connection \(\nabla\) that the manifold \(\mathcal{M}\) is endowed with, and by \(P^i_j\) the components of the parallel transport map \(\mathcal{P}\). Given a vector \(w_y = w^i_y \partial_i\) to be transported, the local expression of the principal pushforward map in \((x, y, w_y) \in \mathcal{M} \times T \mathcal{M}\) reads

\[
(d_x \mathcal{P}^{y \rightarrow x})_{x,y,w_y} = \left( \frac{\partial P^k_j}{\partial x^i} + P^r_j \Gamma_{r i}^k \right) w^i_y d^j_x \otimes \partial_k\big|_x.
\]  

(24)

**Proof.** By definition, it holds that \(d^i_x(\partial_j\big|_x) = \delta^i_j\), \(\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k\) (Einstein summation convention assumed) and \(\mathcal{P}^{y \rightarrow x} = P^i_j d^j_y \otimes \partial_i\big|_x\). Notice that the components of the principal pushforward map are functions of the points \(x, y \in \mathcal{M}\), namely \(P^i_j = P^i_j(x, y)\).

Consider two smooth curves \(x(t) \in \mathcal{M}\) and \(y(t) \in \mathcal{M}\) and define the vector field \(v_x(t) := \mathcal{P}^{y \rightarrow x}(t)(w_y(t)) \in T_x x(t)\). In local coordinates \(v_x = P^i_j w^i_y \partial_i\big|_x\). Decompose the velocity fields of the two curves as \(\dot{x} = \dot{x}^i \partial_i\big|_x\) and \(\dot{y} = \dot{y}^i \partial_i\big|_y\).

The covariant derivative of the vector field \(v_x\) with respect the velocity field of the \(x\)-curve has expression

\[
\nabla_x v_x = \left( \frac{\partial P^k_j}{\partial x^i} x^k + \frac{\partial P^i_j}{\partial y^k} y^k \right) w^j_y \partial_i\big|_x + P^r_j w^r_y \partial_i\big|_x + P^r_j w^r_y (\dot{x}^r \Gamma^k_{r i} \partial_k\big|_x) .
\]  

(25)

The components of the principal pushforward map are those that act on the terms \(\dot{x}^i\), namely, upon renaming indexes,

\[
\frac{\partial P^k_j}{\partial x^i} w^i_y \partial_k\big|_x + P^r_j w^r_y \Gamma^k_{r i} \partial_k\big|_x ,
\]  

(26)

which proves the assertion. \(\square\)

As a side note, it is not difficult to prove that \(P^i_j(x, x) = \delta^i_j\), \(P^i_j(y, y) = \delta^i_j\), which means that the matrix of coefficients \(P^i_j\) is orthogonal, and that \(\Gamma^k_{ij}(z) = \frac{\partial P^k_j(x, y)}{\partial y^r} \big|_{(x, y) = (z, z)}\) for \(z \in \mathcal{M}\), which testifies the intimate relationship between parallel transport and covariant derivation.

In the remainder of the paper, we will restrict our study to the principal pushforward map, that enjoys some interesting algebraic properties, as outlined in the following

**Theorem 3.3.** Denoting by \((d_x \mathcal{P}^{y \rightarrow x})^\dagger_{x,y,w_y} : T_x \mathcal{M} \rightarrow T_y \mathcal{M}\) the adjoint (or dual) of the operator \((d_x \mathcal{P}^{y \rightarrow x})_{x,y,w_y}\) with respect to the Riemannian metric of \(\mathcal{M}\), it holds that

\[
(d_x \mathcal{P}^{y \rightarrow x})^\dagger_{x,y,\mathcal{P}^{x \rightarrow y}(w_x)}(v_x) = 0.
\]  

(27)

---

\(^1\)This content was suggested, and drafted, by one of the anonymous reviewers of the manuscript.
for every $x, y \in \mathbb{M}$ and $v_x \in T_x\mathbb{M}$. Moreover, the ‘diagonal’ part of the linear map $(d_x\mathcal{P}^{y \rightarrow x})_{x,y,w_y}$ is identically zero, namely

$$(d_x\mathcal{P}^{y \rightarrow x})_{x,x,v_x} \equiv 0,$$

for every $x \in \mathbb{M}$ and $v_x \in T_x\mathbb{M}$.

**Proof.** In order to prove the first 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

$$\|v_x(t)\|_t^2 = \|w_y(t)\|_t^2. \quad (29)$$

By taking the derivative of both sides with respect to the parameter $t$, we get

$$\langle v_x, \nabla_x v_x \rangle_x = \langle w_y, \nabla_y w_y \rangle_y. \quad (30)$$

By replacing the expression (15) for the covariant derivative $\nabla_x v_x$ in (30), we get:

$$(v_x, (d_x\mathcal{P}^{y \rightarrow x})_{x,y,w_y}(\dot{x}))_x - \langle v_x, (\mathcal{P}^{y \rightarrow x} \circ (d_x\mathcal{P}^{y \rightarrow x})_{y,x,v_y})(\dot{y}) \rangle_x + \langle \mathcal{P}^{x \rightarrow y}(w_x), \mathcal{P}^{y \rightarrow x}(\nabla_y w_x) \rangle_x = \langle w_x, \nabla_y w_y \rangle_y.$$ 

By the isometry property, the last term on the left-hand side is the same quantity on the right-hand side, therefore, the above property is equivalent to

$$\langle v_x, (d_x\mathcal{P}^{y \rightarrow x})_{x,y,w_y}(\dot{x}) \rangle_x = \langle v_x, (\mathcal{P}^{y \rightarrow x} \circ (d_x\mathcal{P}^{y \rightarrow x})_{y,x,v_y})(\dot{y}) \rangle_x. \quad (31)$$

Let us recall that, given a linear operator $L : \mathcal{V} \rightarrow \mathcal{V}$ on a vector space $\mathcal{V}$ endowed with a positive-definite inner product $\langle \cdot, \cdot \rangle : \mathcal{V}^2 \rightarrow \mathbb{R}$, its adjoint $L^\dagger$ with respect to the metric satisfies $\langle L(v), u \rangle = \langle v, L^\dagger(u) \rangle$ for every $v, v \in \mathcal{V}$. The left-hand side of the equation (31) may thus be rewritten as $(d_x\mathcal{P}^{y \rightarrow x})_{x,y,p,v_y}(v_x, \dot{x})_x$. Since $x, y, \dot{x}$ and $\dot{y}$ are arbitrary, the equation (31) may hold only if the identity (27) is true.

The proof of the second statement follows from the definitions (13), in fact

$$(d_xf)_{x(t),x(t),v_x(t)}(\dot{x}(t)) = \lim_{s \rightarrow 0} \frac{1}{s} \left[ \mathcal{P}^{x(t+s) \rightarrow x(t)}(f(x(t+s), x(t), v_x(t))) - f(x(t), x(t), v_x(t)) \right],$$

$$= \lim_{s \rightarrow 0} \frac{1}{s} \left[ \mathcal{P}^{x(t+s) \rightarrow x(t)}(\mathcal{P}^{x(t) \rightarrow x(t+s)}(v_x(t))) - v_x(t) \right],$$

$$= \lim_{h \rightarrow 0} \frac{1}{h} [v_x(t) - v_x(t)],$$

$$= 0,$$

no matter how we choose $x \in \mathbb{M}$ nor $v_x \in T_x\mathbb{M}$. Since $\dot{x}$ is arbitrary, we get the statement (28).

**Example 8.** Let us recall the following two facts about linear operators:

- **Equivalence of matrix representations:** Given a linear operator $L : \mathcal{V} \rightarrow \mathcal{V}$ on a vector space $\mathcal{V} \subset \mathbb{R}^m$, two of its matrix representations $[L]_1$ and $[L]_2$ are equivalent on the vector space $\mathcal{V}$ if $[L]_1 x = [L]_2 x$ for any $x \in \mathcal{V}$, although the matrices $[L]_1$ and $[L]_2$ may look different.
- **Representation of the adjoint:** If $[L]$ is a matrix representation of a linear operator $L : \mathcal{V} \rightarrow \mathcal{V}$ on a vector space $\mathcal{V} \subset \mathbb{R}^m$ with respect to an orthonormal basis of $\mathcal{V}$, 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 (22) of the operator \((d_x P)_{x,y,w}\) is not unique. For example, if we add to the matrix \(\,(d_x P)_{x,y,w}\) the term \(x^T w_y (1 + x^T y)^{-1} xx^T\), we get the equivalent representation of \((d_x P)_{x,y,w}\):

\[
\|(d_x P)_{x,y,w}\|_1 = -\left(\frac{(x^T w_y)(I + xy^T - xx^T)}{1 + x^T y}\right) + \frac{(x^T w_y)(y + x)y^T}{(1 + x^T y)^2}.
\]

Another equivalent representation is

\[
\|(d_x P)_{x,y,w}\|_2 := \|(d_x P)_{x,y,w}\|_1 - \frac{(w_y^T x)^2}{(1 + x^T y)^2} \frac{[P]_{x,y,w}^T v_x}{\|v_x\|^2},
\]

that serves to exemplify the property (27). In fact, in the present context, it holds that \(\|(d_x P)_{x,y,w}\|_1 = \|(d_x P)_{x,y,w}\|_2\) and (by some lengthy calculations that are omitted here) one may verify that \(\|(d_x P)_{x,y,v_x}\|_2 v_x = 0\), no matter how we choose the points \(x \in S^{n-1}\), \(y \in S^{n-1} \setminus \{-x\}\) and the tangent vector \(v\) in \(T_x S^{n-1}\).

No matter how we choose the matrix representation of the linear operator \((d_x P)_{x,y,w}\), it is quite straightforward to verify directly that \(\|(d_x P)_{x,x,v_x}\| = 0\), as it is assured by the general property (28).

Suppose that the curves \(x(t)\) and \(y(t)\) introduced in the body of the Theorem 3.2 intersect at a point \(p \in M\), as shown in the Figure 3. Since \(v_x = P^{y-x}(w_y)\), when \(x = y = p\) the vector fields \(v_x\) and \(w_y\) align to one another at \(p\), namely \(v_p = w_p \in T_p M\). On the other hand, the two curves are arbitrary and we assume that \(\dot{x}_p \neq \dot{y}_p\).

**Figure 3.** Exemplification of two curves on a curved manifold \(M\) that meet at a point \(p\).

The Theorem 3.2 and the Theorem 3.3 together tell an interesting property of the total derivatives of the vector fields \(v_x\) and \(w_y\) at the point where the two curves, that they are defined on, meet.
Corollary 2. At the point \( p \), \( \nabla_x v_x = \nabla_y w_y \).

Proof. By the Theorem 3.2, at any points \( x, y \in \mathbb{M} \) and for every \( w \in T_y \mathbb{M} \), it holds that
\[
(\nabla_x v_x)_x = (dx \mathcal{P} y \rightarrow x)_{x, y, w} (\dot{x}_p) - \mathcal{P} y \rightarrow x \left[(dx \mathcal{P} y \rightarrow x)_{y, x, \mathcal{P} y \rightarrow x} (\dot{y}_p) - (\nabla_y w_y)_y\right].
\]
Setting \( x = y = p \) gives
\[
(\nabla_x v_x)_p = (dx \mathcal{P} y \rightarrow x)_{p, p, w} (\dot{x}_p) - (dx \mathcal{P} y \rightarrow x)_{p, p, w} (\dot{y}_p) + (\nabla_y w_y)_p,
\]
namely
\[
(\nabla_x v_x)_p - (\nabla_y w_y)_p = (dx \mathcal{P} y \rightarrow x)_{p, p, w} (\dot{x}_p - \dot{y}_p).
\]
By virtue of the Theorem 3.3, the right-hand side is identically zero, from which the conclusion follows.

Example 9. The statement of the Corollary 2, written explicitly for the unit hypersphere, leads to non-trivial constraints on the na"ive derivatives \( (v'_x)_p, (w'_y)_p \in \mathbb{A} \). By using the formula (8) for the covariant derivative, the statement of the Corollary reads
\[
(v'_x)_p + pu^T p \dot{x}_p = (w'_y)_p + pw^T p \dot{y}_p.
\]
Since \( u_p = w_p \), the above equation may be rewritten, for example, as
\[
(v'_x)_p - (w'_y)_p = v^T p (\dot{y}_p - \dot{x}_p).
\]
If \( \dot{y}_p = \dot{x}_p \), then trivially \( (v'_x)_p = (w'_y)_p \), otherwise, they differ by the amount \( v^T p (\dot{y}_p - \dot{x}_p) \) along the radial direction \( p \).

4. Cases study: Synchronization of first-order and second-order non-linear oscillators on manifolds. Manifold calculus is the principal theoretical tool to analyze control problems for systems that need to obey holonomic (as well as non-holonomic) constrains. Exemplary early contributions are the studies on generalized Dubins problem [6, 7, 8] and on the rolling problem in arbitrary dimensions [22]. Recent contribution to this field are the paper [3] that adapts the multivariate optimal control theory to a Riemannian setting, and the paper [32] that proposes a modified integral controller to treat actuation biases in nonholonomic systems.

The present section studies two cases of interest, namely, the problem of synchronizing two non-linear oscillators of the first order and of the second order (Duffing-type). In the first case, the systems to synchronize are of the first order and the principal pushforward map is not required to analyze the error system. In contrast, in the second case, the twin systems to synchronize are of the second order and it is necessary to make use of the principal pushforward map to define and compute a synchronizing control field, as well as to characterize the error system dynamics.

4.1. Synchronization of first order oscillators on a Riemannian manifold. The goal of synchronization is to analyze control problems for systems that need to obey holonomic (as well as non-holonomic) constrains. Exemplary early contributions are the studies on generalized Dubins problem [6, 7, 8] and on the rolling problem in arbitrary dimensions [22]. Recent contribution to this field are the paper [3] that adapts the multivariate optimal control theory to a Riemannian setting, and the paper [32] that proposes a modified integral controller to treat actuation biases in nonholonomic systems.

The present section studies two cases of interest, namely, the problem of synchronizing two non-linear oscillators of the first order and of the second order (Duffing-type). In the first case, the systems to synchronize are of the first order and the principal pushforward map is not required to analyze the error system. In contrast, in the second case, the twin systems to synchronize are of the second order and it is necessary to make use of the principal pushforward map to define and compute a synchronizing control field, as well as to characterize the error system dynamics.

4.1. Synchronization of first order oscillators on a Riemannian manifold. The goal of synchronization is to analyze control problems for systems that need to obey holonomic (as well as non-holonomic) constrains. Exemplary early contributions are the studies on generalized Dubins problem [6, 7, 8] and on the rolling problem in arbitrary dimensions [22]. Recent contribution to this field are the paper [3] that adapts the multivariate optimal control theory to a Riemannian setting, and the paper [32] that proposes a modified integral controller to treat actuation biases in nonholonomic systems.

The present section studies two cases of interest, namely, the problem of synchronizing two non-linear oscillators of the first order and of the second order (Duffing-type). In the first case, the systems to synchronize are of the first order and the principal pushforward map is not required to analyze the error system. In contrast, in the second case, the twin systems to synchronize are of the second order and it is necessary to make use of the principal pushforward map to define and compute a synchronizing control field, as well as to characterize the error system dynamics.

4.1. Synchronization of first order oscillators on a Riemannian manifold. The goal of synchronization is to analyze control problems for systems that need to obey holonomic (as well as non-holonomic) constrains. Exemplary early contributions are the studies on generalized Dubins problem [6, 7, 8] and on the rolling problem in arbitrary dimensions [22]. Recent contribution to this field are the paper [3] that adapts the multivariate optimal control theory to a Riemannian setting, and the paper [32] that proposes a modified integral controller to treat actuation biases in nonholonomic systems.

The present section studies two cases of interest, namely, the problem of synchronizing two non-linear oscillators of the first order and of the second order (Duffing-type). In the first case, the systems to synchronize are of the first order and the principal pushforward map is not required to analyze the error system. In contrast, in the second case, the twin systems to synchronize are of the second order and it is necessary to make use of the principal pushforward map to define and compute a synchronizing control field, as well as to characterize the error system dynamics.
systems on a manifold $\mathbb{M}$. The leader and the follower are described by the equations

$$\begin{align*}
\dot{r} &= \sigma(r), \\
\dot{x} &= \sigma(x) + u,
\end{align*}$$

(38)

where $\sigma : \mathbb{M} \to T\mathbb{M}$ is the state-transition function and $u \in T\mathbb{M}$ is a tangent control field, namely $u(t) \in T_{x(t)}\mathbb{M}$.

According to the error-based dynamical system theory summarized in [15, 24], we define two error fields $(e, \varepsilon) \in (T\mathbb{M})^2$ such that

$$\begin{align*}
e &= -\log_x r \quad \text{(synchronization error)}, \\
\varepsilon &= \dot{x} - \mathcal{P}_{r \to x}(\dot{r}).
\end{align*}$$

(39)

By definition of the synchronization error and by the system-kinetic equations (38), it holds that:

$$\varepsilon = \sigma(x) + u - \mathcal{P}_{r \to x}(\sigma(r)).$$

(40)

The choice of control field proposed in [13] to synchronize the twin systems is

$$u := \mathcal{P}_{r \to x}(\sigma(r)) - \sigma(x) - ce,$$

(41)

with $c > 0$ being a constant coefficient termed \textit{communication strength}.

The resulting error dynamics is governed by the equation

$$\varepsilon + ce = 0.$$

(42)

The above first-order error system is characterized by the following

\textbf{Theorem 4.1.} \textit{The synchronization error converges exponentially fast to the state zero, hence, the control field (41) asymptotically synchronizes the follower to the leader.}

\textit{Proof.} The convergence of the synchronization error to zero is proven through the Lyapunov function

$$\mathcal{W}(x, r) := \frac{1}{2}d^2(x, r) = \frac{1}{2}\|e\|^2_x,$$

(43)

where $d : \mathbb{M}^2 \to \mathbb{R}$ denotes a Riemannian distance and $\log : \mathbb{M}^2 \to T\mathbb{M}$ denotes a manifold logarithmic map (namely, the inverse of the manifold exponential map). Deriving the function $\mathcal{W}$ with respect to the time gives:

$$\frac{d\mathcal{W}}{dt} = (\langle -\log_x r, \dot{x} \rangle_x + \langle -\log_x x, \dot{r} \rangle_r).$$

(44)

Recall from Subsection 2.1 that parallel transport is an isometry, therefore:

$$\langle \log_x x, \dot{r} \rangle_r = \langle \mathcal{P}_{r \to x} \log_x x, \mathcal{P}_{r \to x} \dot{r} \rangle_x.$$

(45)

Moreover, since parallel transport moves a tangent vector along a geodesic connecting the starting and the end points, it holds that $\mathcal{P}_{r \to x} \log_x x = -\log_x r$ [13]. Therefore

$$\mathcal{W} = \langle e, \dot{x} - \mathcal{P}_{r \to x} \dot{r} \rangle_x$$

$$= \langle e, \varepsilon \rangle_x,$$

$$= \langle e, -ce \rangle_x,$$

$$= -c \langle e, e \rangle_x.$$

(46)

The function $\mathcal{W}(x, r) \geq 0$ and, since the communication strength is positive, its time-derivative $\dot{\mathcal{W}} = -c\|e\|^2_x \leq 0$, hence $\mathcal{W}$ is a Lyapunov function for the error system (42). From the equations (46), it follows that

$$\dot{\mathcal{W}} = -2c\mathcal{W},$$

(47)
hence the synchronization error tends towards zero exponentially fast.

As anticipated, the study of a first-order error system, corresponding to a first-order system dynamics, does not require a principal pushforward map. In contrast, the following case of study, that concerns a second-order Duffing-type dynamical system, will need to make use of a principal pushforward map to compute a synchronizing control field as well as to study the associated error system dynamics.

4.2. Synchronization of Duffing oscillators on a Riemannian manifold. As a further case study, we consider the synchronization of two identical Duffing-type oscillators on a Riemannian manifold \( M \). According to an extension of classical nonlinear oscillators introduced in [12], a Duffing-type oscillator on a Riemannian manifold is described by the equations

\[
\begin{align*}
\dot{x} &= v, \\
\nabla \dot{x} v &= -\alpha v + \nabla_x V,
\end{align*}
\]  

(48)

where \( \alpha > 0 \) is a damping constant and \( \nabla_x V \) denotes the Riemannian gradient of the potential \( V : M \to \mathbb{R} \) defined as

\[
V(x) = \frac{1}{2} \beta d^2(x, x_0) + \frac{1}{4} \gamma d^4(x, x_0),
\]  

(49)

where \( \beta, \gamma \in \mathbb{R} \) are arbitrary constants that define the shape of the potential function, and \( x_0 \in M \) denotes a fixed reference point around which the oscillations take place.

For the sake of notational compactness, let us introduce the state-transition function

\[
\sigma(x, v) := -\alpha v + \nabla_x V = -\alpha v - (\beta + \gamma d^2(x, x_0)) \log d(x, x_0).
\]  

(50)

The equations for the leader system, whose evolution is described by the state-variable \( \mathbf{r} \), and the follower system, whose evolution in time is described by the state-variable \( x \), may be written as

\[
\begin{align*}
\nabla \dot{x} \mathbf{r} &= \sigma(r, \dot{r}), \\
\nabla \dot{x} \mathbf{x} &= \sigma(x, \dot{x}) + u,
\end{align*}
\]  

(51)

where \( u \in TM \) is a control field, namely \( u(t) \in T_{\mathbf{x}(t)} M \).

Let us define again an error-fields pair \( (e, \varepsilon) \in (TM)^2 \) as in (39). Taking the covariant derivative of the error field \( \varepsilon \) leads to:

\[
\nabla \varepsilon = \nabla \dot{x} - \nabla \dot{x} \mathcal{D}^{\mathbf{r} \to \mathbf{r}}(\dot{r}).
\]  

(52)

According to the Theorem 3.2, it holds that

\[
\nabla \dot{x} \mathcal{D}^{\mathbf{r} \to \mathbf{r}}(\dot{r}) = (d_x \mathcal{D}^{\mathbf{r} \to \mathbf{r}})_{\mathbf{x}, \mathbf{r}, \dot{r}}(\dot{x}) - \mathcal{D}^{\mathbf{r} \to \mathbf{r}}[d_x \mathcal{D}^{\mathbf{r} \to \mathbf{r}}]_{\mathbf{x}, \mathbf{r}, \mathcal{D}^{\mathbf{r} \to \mathbf{r}}(\dot{r})}(\dot{r}) - \nabla \dot{r}^r,
\]  

(53)

Plugging the equation (53) and the equations (51) into the equation (52), yields

\[
\nabla \varepsilon = \sigma(x, \dot{x}) - \mathcal{D}^{\mathbf{r} \to \mathbf{r}}(\sigma(r, \dot{r})) + \mathcal{D}^{\mathbf{r} \to \mathbf{r}}[(d_x \mathcal{D}^{\mathbf{r} \to \mathbf{r}})]_{\mathbf{x}, \mathbf{r}, \mathcal{D}^{\mathbf{r} \to \mathbf{r}}}(\dot{r}) + u.
\]  

(54)

On the basis of a general theory to design non-linear synchronizing controls recalled in [23], a strategy to choose the control field \( u \) is to make it cancel the terms that do not depend on the errors \( e, \varepsilon \) and to add terms in the error that
guarantee a desired performance. In the present case study, all terms on the right-hand side of the equation (54) will be canceled out and the added terms will be taken as in [24], namely:

\[
\begin{align*}
    u &:= \mathcal{P}^{r \to x}(\sigma(r, \dot{r})) - \sigma(x, \dot{x}) + (d_x \mathcal{P}^{r \to x})_{x,r,\dot{r}}(\dot{x}) \\
    &\quad - \mathcal{P}^{r \to x}((d_x \mathcal{P}^{r \to x})_{r,x,\mathcal{P}^{r \to x}(\dot{r})}) - \kappa_P \epsilon - \kappa_D \varepsilon,
\end{align*}
\]  

with \(\kappa_P, \kappa_D\) constants, with \(\kappa_D > 0\) termed derivative coefficient. Notice that, in order to compute the value of the control field, it is necessary to employ the principal pushforward map repeatedly.

The resulting error dynamics is governed by the equation

\[
\nabla_x \varepsilon + \kappa_P \epsilon + \kappa_D \varepsilon = 0.
\]  

The above second-order error system is characterized by the following

**Theorem 4.2.** The error system (56) converges asymptotically to the state zero, hence, the control field (55) asymptotically synchronizes the follower to the leader oscillator.

**Proof.** The convergence of the synchronization error to zero is proven through the Lyapunov function

\[
\mathcal{W}(x, r) := \frac{1}{2} \langle \varepsilon, \varepsilon \rangle_x + \kappa_P d^2(x, r).
\]  

Deriving the function \(\mathcal{W}\) with respect to the time gives:

\[
\frac{d \mathcal{W}}{dt} = \frac{1}{2} \langle \nabla_x \varepsilon, \varepsilon \rangle_x + \frac{1}{2} \langle \epsilon, \nabla_x \varepsilon \rangle_x + \kappa_P \langle (-\log_x r, \dot{x})_x + (-\log_r x, \dot{r})_r \rangle.
\]  

Recalling again the properties of parallel transport leads to

\[
\begin{align*}
    \frac{d \mathcal{W}}{dt} &\leq \langle \nabla_x \varepsilon, \varepsilon \rangle_x + \kappa_P \langle \epsilon, \dot{\varepsilon} \rangle_x \\
    &\quad + \langle \kappa_D \varepsilon, \varepsilon \rangle_x + \kappa_P \langle \epsilon, \varepsilon \rangle_x \\
    &\quad + \langle -\kappa_D \epsilon - \kappa_P \epsilon + \kappa_P \epsilon, \varepsilon \rangle_x \\
    &\quad = -\kappa_D \langle \varepsilon, \varepsilon \rangle_x.
\end{align*}
\]  

Since the derivative coefficient is positive, the function \(\mathcal{W}(x, r) := \frac{1}{2} \| \varepsilon \|_x^2 + \kappa_P d^2(x, r) \geq 0\) and its time-derivative \(\dot{\mathcal{W}} = -\kappa_D \| \varepsilon \|_x^2 \leq 0\), hence \(\mathcal{W}\) is a Lyapunov function for the error system (56).

Notice that the above Theorem is independent of the chosen potential function \(\mathcal{V}\), hence it applies to any Duffing-type oscillators pairs. Also, notice that, in contrast to the proof presented in [24], there’s no additional compatibility condition to set to study the error system dynamics.

As a side note, the proof presented here differs substantially from the proof presented in the paper [24], which assumes that \(\varepsilon = \nabla_x e\) rather than defining \(\varepsilon\) and \(e\) separately. Informally speaking, we based our proof on the observation that there is no need to require \(\varepsilon\) and \(e\) to be related by a covariant derivation.
5. Relationship between the principal pushforward map and the curvature endomorphism. A parallel transport operator does not behave trivially under composition. For example, given three points \( x, y, z \in \mathbb{M} \), in general \( \mathcal{P}^{y \rightarrow z} \circ \mathcal{P}^{z \rightarrow y} \neq \mathcal{P}^{y \rightarrow z} \), unless these points belong to the same geodesic line. This sort of observations lead to a whole theory termed holonomy which, in turn, is closely related to the local curvature properties of a state manifold connection [29].

Let us recall that to any point \( x \) of a connected manifold \( \mathbb{M} \) endowed with a connection \( \nabla \), we may associate an holonomy group \( \text{Hol}_x(\nabla) \subset \text{Aut}(T_x\mathbb{M}) \) formed by all the parallel transports along a piecewise smooth closed loop originating and ending at \( x \) [29]. For example, given any three distinct points \( x, y, z \in \mathbb{M} \) and a tangent vector \( w_y \in T_y\mathbb{M} \), we may transport \( w_y \) from \( y \) to \( z \) by means of the parallel transport operator \( \mathcal{P}^{y \rightarrow z} \), then we may transport the resulting tangent vector from the point \( z \) to the point \( x \) along the geodesic line connecting them by means of the parallel transport operator \( \mathcal{P}^{z \rightarrow x} \), and then we may transport the resulting tangent vector from the point \( x \) back to the point \( y \) along the geodesic line connecting them by means of the parallel transport operator \( \mathcal{P}^{x \rightarrow y} \); in this case, the closed loop is referred to as geodesic triangle. The resulting vector

\[
\bar{w}_y := (\mathcal{P}^{x \rightarrow y} \circ \mathcal{P}^{z \rightarrow x} \circ \mathcal{P}^{y \rightarrow z})(w_y)
\]

will, in general, differ from the original vector \( w_y \), because of the curvature of the manifold, which gives rise to the notion of holonomy. The compound linear map \( \mathcal{H}_y(z, x) := \mathcal{P}^{x \rightarrow y} \circ \mathcal{P}^{z \rightarrow x} \circ \mathcal{P}^{y \rightarrow z} \) is an element of the Holonomy group, namely, \( \mathcal{H}_y(z, x) \in \text{Hol}_y(\nabla) \). Since each element of the parallel transport chain is an isometry, the map \( \mathcal{H}_y(z, x) \) is an isometry as well, therefore \( \|\bar{w}_y\|_y = \|w_y\|_y \). (In fact, if one chooses an orthogonal basis for \( T_y\mathbb{M} \), the holonomy group \( \text{Hol}_y(\nabla) \) is a subgroup of the orthogonal group \( \text{O}(m) \), with \( m \) being the dimension of the manifold \( \mathbb{M} \).)

Let us recall that, on a Riemannian manifold \( \mathbb{M} \) endowed with a connection \( \nabla \), we may define a Riemannian curvature operator \( \mathcal{R} : (TM)^3 \rightarrow TM \) by

\[
\mathcal{R}_x(u_x, v_x)w_x := \nabla_{u_x} \nabla_{v_x} w_x - \nabla_{v_x} \nabla_{u_x} w_x - [u_x, v_x]_x w_x,
\]

which essentially measures the non-commutativity of the covariant derivative due to the curved nature of the manifold\(^2\). In the above definition, \( x \in \mathbb{M}, x \mapsto u_x, v_x, w_x \) are tangent vector fields and \([\cdot, \cdot]_x \) denotes the Lie bracket of vector fields in \( TM \).

The linear transformation \( \mathcal{R}_x(u_x, v_x) \in \text{End}(T_x\mathbb{M}) \) is referred to as the curvature endomorphism. Two important properties of the curvature endomorphism are

- **Bilinearity**: For every \( u_x, v_x \in T_x\mathbb{M} \) and \( \alpha \in \mathbb{R} \), it holds that \( \mathcal{R}_x(\alpha u_x, v_x) = \alpha \mathcal{R}_x(u_x, v_x) \).

- **Skew-symmetry**: For every \( u_x, v_x \in T_x\mathbb{M} \), it holds that \( \mathcal{R}_x(u_x, v_x) = -\mathcal{R}_x(v_x, u_x) \).

There exists a close relationship between the holonomy group and the curvature endomorphisms associated to a connection.

\(^2\)Indeed, a compact definition of the curvature endomorphism that emphasizes its feature of being a measure of non-commutativity of the covariant derivative is \( \mathcal{R}(u, v) := [\nabla_u, \nabla_v] - \nabla_{[u, v]} \). To put the Riemannian curvature in context, it is worth recalling that the components of the curvature form a \((1,3)\)-tensor. Due to the symmetry of the tensor, contracting any two indexes produces a null tensor, except for one contraction that leads to the Ricci \((0,2)\)-tensor, which plays a prominent role in the Einsteinian description of the space-time [21]. A further contraction leads to the scalar curvature.
An interesting relationship between the principal pushforward map \((d_x \mathcal{S}^{y \rightarrow x})_{x,y,w_y}\) and the manifold holonomy comes from the definition \((13)\). In fact, this definition may be rewritten equivalently as

\[
(d_x \mathcal{S}^{y \rightarrow x})_{x,y,w_y}(\dot{x}(t)) = \lim_{s \to 0} \frac{1}{s} \left[ (\mathcal{S}^{x(t+s) \rightarrow x}(t) \circ \mathcal{S}^{y(t) \rightarrow x(t+s)})(w_y(t)) - \mathcal{S}^{y(t) \rightarrow x(t)}(w_y(t)) \right] = \lim_{s \to 0} \frac{1}{s} \left[ (\mathcal{S}^{x(t) \rightarrow y(t)} \circ \mathcal{S}^{x(t+s) \rightarrow y(t)} \circ \mathcal{S}^{y(t) \rightarrow x(t)}) (w_y(t)) - w_y(t) \right] = \mathcal{S}^{y(t) \rightarrow x(t)} \left\{ \lim_{s \to 0} \frac{1}{s} \left[ (\mathcal{S}^{x(t) \rightarrow y(t)} \circ \mathcal{S}^{x(t+s) \rightarrow x(t)} \circ \mathcal{S}^{y(t) \rightarrow x(t+s)}) (w_y(t)) - w_y(t) \right] \right\}.
\]

A close look at the last line of the above relationship reveals that the argument of the limit in the right-hand side involves a comparison between the vector \(w_y\) and the result of its transformation by the element \(\mathcal{S}^{x(t) \rightarrow y(t)} \circ \mathcal{S}^{x(t+s) \rightarrow x(t)} \circ \mathcal{S}^{y(t) \rightarrow x(t+s)}\) of the holonomy group \(\text{Hol}_y(\nabla)\) and

\[
(d_x \mathcal{S}^{y \rightarrow x})_{x,y,w_y}(\dot{x}) = \mathcal{S}^{y \rightarrow x} \left\{ \lim_{s \to 0} \frac{1}{s} \left[ \mathcal{H}_y(x(\cdot + s), x) - \text{id}_{y} \right] w_y \right\},
\]

in shortened notation.

In order to elucidate the relationship between the principal pushforward map and the curvature endomorphism, it pays to recall a fundamental result from the holonomy theory (along with its interesting proof). As a reference, we point the reader to the paper \([27]\), that generalizes the following result to arbitrary fiber bundles.

**Lemma 5.1.** Let \(H : \mathbb{I} \times [0, 1] \to \mathcal{M}\) denote a geodesic homotopy, namely, for any fixed \(s \in \mathbb{I}, \theta \mapsto H(s, \theta)\) traces a geodesic curve on \(\mathcal{M}\) connecting the point \(H(s, 0)\) to the point \(H(s, 1)\). Let \(\sigma : \mathbb{I} \times [0, 1] \to T\mathcal{M}\) denote a smooth vector field along the homotopy, namely, \(\sigma(s, \theta) \in T_{H(s, \theta)}\mathcal{M}\). Let \(\nabla\) denote a Riemannian (i.e., torsion free) connection. Assume that the vector field \(\sigma\) is uniform along the homotopy, namely, that

\[
(\nabla_{\partial_s H})\sigma(s, 0) = 0 \quad \text{and} \quad (\nabla_{\partial_\theta H})\sigma(s, \theta) = 0
\]

for every value of \(s \in \mathbb{I}\) and for every value of \(\theta \in [0, 1]\). Then, it holds that

\[
(\nabla_{\partial_s H})\sigma(s, 1) = \left( \int_0^1 \mathcal{H}^{H(s, \theta) \rightarrow H(s, 1)} \circ \mathcal{H}^{H(s, 0)}(\partial_s H, \partial_\theta H) \circ \mathcal{H}^{H(s, 0) \rightarrow H(s, \theta)} \, d\theta \right) \sigma(s, 1).
\]

**Proof.** The quantities involved in the Lemma and in the present proof are represented in the Figure 4. The partial derivative \(\partial_s H \in T_{H(s, \theta)}\mathcal{M}\) denotes the velocity vector associated to the geodesic curve \(\theta \mapsto H(s, \theta)\) corresponding to a fixed value of the parameter \(s\), while the partial derivative \(\partial_\theta H \in T_{H(s, \theta)}\mathcal{M}\) denotes a transverse tangent direction. By the definition \((60)\) of Riemannian curvature, we have that

\[
\nabla_{\partial_s H} \nabla_{\partial_\theta H} \sigma = \nabla_{\partial_\theta H} \nabla_{\partial_s H} \sigma + \nabla_{[\partial_s H, \partial_\theta H]} \sigma + \mathcal{R}_H(\partial_s H, \partial_\theta H) \sigma.
\]

Since \(\nabla\) is torsion free, \([\partial_s H, \partial_\theta H] = 0\). Moreover, by hypothesis, the vector field \(\sigma\) is parallel along the homotopy, namely, \(\nabla_{\partial_s H} \sigma = 0\). Therefore, the expression \((64)\) simplifies to

\[
(\nabla_{\partial_\theta H} \nabla_{\partial_s H})\sigma(s, \theta) = \mathcal{R}_H(s, \theta)(\partial_s H, \partial_\theta H) \sigma(s, \theta).
\]

If \(m\) denotes the dimension of the manifold \(\mathcal{M}\) and \((\sigma_1(s, 0), \sigma_2(s, 0), \ldots, \sigma_m(s, 0))\) denotes a basis of the tangent space \(T_{H(s, 0)}\mathcal{M}\), one
may construct a basis of each tangent space $T_{H(s, \theta)}M$ by parallelly-transporting each basis vector from $H(s, 0)$ to $H(s, \theta)$, namely by setting

$$\sigma_i(s, \theta) := \mathcal{P}^{H(s, 0) \to H(s, \theta)}(\sigma_i(s, 0)), \quad i = 1, 2, \ldots, m.$$  \hfill (66)

With respect to this basis, the vector field $\nabla_{\partial_s H} \sigma$ may be decomposed as

$$\nabla_{\partial_s H} \sigma(s, \theta) = \sum_i q^i(s, \theta) \sigma_i(s, \theta),$$  \hfill (67)

where the smooth function $q := (q^1, q^2, \ldots, q^m) : \mathbb{I} \times [0, 1] \to \mathbb{R}^m$ is termed principal part of the vector field $\nabla_{\partial_s H} \sigma$.

Since the vector fields $\sigma_i$ are parallel along the geodesic lines constituting the homotopic net, it holds that $\nabla_{\partial_s H} \sigma_i = 0$, therefore, by the Leibniz rule for the covariant derivative, one gets that

$$\nabla_{\partial_\theta H} \nabla_{\partial_s H} \sigma(s, \theta) = \sum_i (\partial_\theta q^i)(s, \theta) \sigma_i(s, \theta).$$  \hfill (68)

Applying the parallel transport operator $\mathcal{P}^{H(s,\theta) \to H(s,1)}$ to both sides gives

$$\mathcal{P}^{H(s,\theta) \to H(s,1)}((\nabla_{\partial_s H} \nabla_{\partial_\theta H} \sigma)(s, \theta)) = \sum_i (\partial_\theta q^i)(s, \theta) \mathcal{P}^{H(s,\theta) \to H(s,1)}(\sigma_i(s, \theta)) = \sum_i (\partial_\theta q^i)(s, \theta) \sigma_i(s, 1).$$  \hfill (69)

Replacing the double covariant derivative $\nabla_{\partial_s H} \nabla_{\partial_\theta H} \sigma$ with the right-hand side of the expression (65) gives

$$\mathcal{P}^{H(s,\theta) \to H(s,1)}(\mathcal{R}_{H(s,\theta)}(\partial_\theta H, \partial_s H) \sigma(s, \theta)) = \sum_i (\partial_\theta q^i)(s, \theta) \sigma_i(s, 1),$$  \hfill (70)

namely,

$$\sum_i (\partial_\theta q^i)(s, \theta) \sigma_i(s, 1) = \left[ \mathcal{P}^{H(s,\theta) \to H(s,1)} \circ \mathcal{R}_{H(s,\theta)}(\partial_\theta H, \partial_s H) \circ \mathcal{P}^{H(s,1) \to H(s,\theta)} \right] \sigma(s, 1).$$  \hfill (71)
Now, the integral of the left-hand side with respect to the parameter $\theta$ reads
\[
\int_0^1 \sum_i (\partial_0 q^i)(s, \theta) \sigma_i(s, 1) d\theta
\]
\[
= \sum_i q^i(s, 1) \sigma_i(s, 1) - \sum_i q^i(s, 0) \sigma_i(s, 1)
\]
\[
= (\nabla_{\partial_0} H \sigma)(s, 1) - \sum_i q^i(s, 0) \rho^{H(s,0)\rightarrow H(s,1)}(\sum_i q^i(s, 0) \sigma_i(s, 0))
\]
\[
= (\nabla_{\partial_0} H \sigma)(s, 1) - \rho^{H(s,0)\rightarrow H(s,1)}((\nabla_{\partial_0} H \sigma)(s, 0)).
\]
By hypothesis, the last term is identically zero, therefore
\[
\int_0^1 \sum_i (\partial_0 q^i)(s, \theta) \sigma_i(s, 1) d\theta = (\nabla_{\partial_0} H \sigma)(s, 1),
\] (72)
while the integral of the right-hand side of the expression (71) with respect to the parameter $\theta$ coincides with the right-hand side of the expression (63), which proves the assertion.

On the basis of the formula (63), it is possible to express the principal pushforward map in terms of the Riemannian curvature endomorphism. This relationship is elucidated by the following result.

**Theorem 5.2.** Let $\exp : T\mathcal{M} \rightarrow \mathcal{M}$ denote the exponential map associated with the connection $\nabla$ of the manifold $\mathcal{M}$ and let $\log : \mathcal{M}^2 \rightarrow T\mathcal{M}$ denote its inverse, namely, the logarithmic map associated to the connection. The principal pushforward map may be expressed in closed form as

\[
\left\{
\begin{array}{l}
(d_x \rho^{y\rightarrow x})_{y,w}(\dot{x}) = (\mathcal{A}^0 \circ \rho^{y\rightarrow x})(w_y),
\end{array}
\right.
\]
where
\[
\mathcal{A}^0 := \int_0^1 \rho^{\exp_y(\theta \log_y x) \rightarrow x} \circ \mathcal{A}^0_{x,y,\dot{x}}(\theta) \circ \rho^{\exp_x(\theta \log_x y) \rightarrow y} \theta d\theta,
\]
\[
\mathcal{A}^0_{x,y,\dot{x}}(\theta) := \rho^{\exp_y(\theta \log_y x)}((d\exp)_y(\theta \log_y x)(\log_y x), (d\exp)_y(\theta \log_y x)(\log_y x)).
\] (73)

**Proof.** In the Lemma 5.1, choose the homotopy $H : \mathbb{I} \times [0, 1] \rightarrow \mathcal{M}$ as
\[
H(s, \theta) := \exp_y(\theta \log_y \exp_x(\dot{x})),
\] (74)
where $\mathbb{I} \ni 0$, and the vector field $\sigma : \mathbb{I} \times [0, 1] \rightarrow T\mathcal{M}$ along said homotopy as
\[
\sigma(s, \theta) := \rho^{y\rightarrow H(s,\theta)}(w_y).
\] (75)

Apparently, the homotopy (74) is smooth and the vector field (75) satisfies both conditions (62), therefore, it holds that

\[
\left\{
\begin{array}{l}
(\nabla_{\partial_0} H \sigma)(s, 1) = \mathcal{A}(\sigma)(s, 1),
\end{array}
\right.
\]
\[
\mathcal{A}(s) := \int_0^1 \rho^{H(s,\theta)\rightarrow H(s,1)}(\partial_0 H, \partial_s H) \circ \rho^{H(s,1)\rightarrow H(s,\theta)} d\theta,
\] (76)
where $\mathcal{A}(s) \in \text{End}(T_{H(s,1)}\mathcal{M})$ for every $s \in \mathbb{I}$.

The relationship between the covariant derivative $(\nabla_{\partial_0} H \sigma)(s, 1)$ and the principal pushforward map is given by the observation that
\[
(d_x \rho^{y\rightarrow x})_{x(t),y(t),w_{y(t)}}(\dot{x}(t)) = \lim_{s \to 0} \frac{1}{s} \left[ (\rho^{x(t+s)\rightarrow x(t)}(\sigma(s, 1) - \sigma(0, 1))
\right]
\]
\[
(\nabla_{\partial_0} H \sigma)(0, 1).
\] (77)
Therefore,
\[
\left\{ \begin{array}{l}
(d_x \mathcal{G}^{y \rightarrow z})_{x(t),y(t),w_y(t)}(\dot{x}(t)) = \mathcal{A}(0)(\sigma(0,1)),
\mathcal{A}(0) = \int_0^1 \mathcal{G}^{H(0,\theta) \rightarrow H(0,1)} \circ \mathcal{H}(0,\theta) \circ \mathcal{H}(0,1) \rightarrow H(0,0) d\theta,
\end{array} \right.
\]
where
\[
H(0,1) = x(t),
\sigma(0,1) = \mathcal{G}^{y(t) \rightarrow x(t)}(w_y(t)),
H(0,\theta) = \exp_{y(t)}(\theta \log_{y(t)} x(t)).
\tag{78}
\]

The partial derivative of the homotopy (74) with respect to the parameter \(\theta\) reads
\[
\partial_\theta H = (d \exp)_{y,\theta \log_y x \exp_x(s \dot{x})}(\log_y x).
\tag{82}
\]

The partial derivative of the homotopy (74) with respect to the parameter \(s\) reads
\[
\partial_s H = (d \exp)_{y,\theta \log_y x \exp_x(s \dot{x})}((d \exp)_{x,s \dot{x}}(\dot{x})),
\tag{83}
\]
where
\[
(d_x \log)_{y,x(t)}(\dot{x}(t)) := \lim_{s \to 0} \frac{1}{s} [\log_y x(t + s) - \log_y x(t)].
\tag{83}
\]

Therefore, said partial derivative at \(s = 0\) reads
\[
\partial_s H |_{s=0} = (d \exp)_{y,\theta \log_y x \exp_x(0 \dot{x})}(0 \dot{x}).
\tag{84}
\]

Since \((d \exp)_{x,0} = \text{id}_x\) and the pushforward map \(d \exp\) is linear, the above formula may be simplified to
\[
\partial_s H |_{s=0} = \theta((d \exp)_{y,\theta \log_y x \circ (d_x \log)}_{y,x}(\dot{x})).
\tag{84}
\]

The above expression was written so as to emphasize the fact that the partial derivative \(\partial_s H |_{s=0}\) is linear in \(\dot{x}\).

In conclusion, replacing the formulas (82) and (84) into the expression (78), we get:
\[
\left\{ \begin{array}{l}
(d_x \mathcal{G}^{y \rightarrow z})_{x,y,w_y}(\dot{x}) = (\mathcal{A}(0) \circ \mathcal{G}^{y \rightarrow z})(w_y),
\mathcal{A}(0) = \int_0^1 \mathcal{G}^{\exp_y(\theta \log_y x) \rightarrow x} \circ \mathcal{A}_{x,y,x}(\theta) \circ \mathcal{G}^{\exp_x(\theta \log_y x)} d\theta,
\mathcal{A}_{x,y,x}(\theta) := \mathcal{A}_{x,y,x}(\theta \log_y x)((d \exp)_{y,\theta \log_y x \circ (d_x \log)}_{y,x}(\dot{x})).
\end{array} \right.
\tag{85}
\]

Since the function \(\mathcal{A}_{x,y,x}(\theta)\) is based on the curvature endomorphism, it is linear in the variable \(\dot{x}\) and \(\mathcal{A}_{x,y,x}(\theta) = \theta \mathcal{A}_{x,y,x}(\theta \log_y x)((d \exp)_{y,\theta \log_y x \circ (d_x \log)}_{y,x}(\dot{x}))\), which justifies the assertion (73). Note that the integral in the second equation of (73) is well defined since the integrand is an endomorphism in \(\text{End}(T_x \mathbb{M})\).

**Example 10.** For the unit hypersphere \(S^{n-1}\) endowed with the canonical metrics inherited by the Euclidean inner product, the curvature endomorphism reads
\[
\mathcal{A}(u,v)w := (v^T w)u - (u^T w)v.
\tag{86}
\]
for every $u, v, w \in T_x \mathbb{S}^{n-1}$ and $x \in \mathbb{S}^{n-1}$. It is immediate to verify that such vector belongs to the subspace $T_x \mathbb{S}^{n-1}$ (since it is a linear combination of the tangent vectors $u$ and $v$). Note that, in this circumstance, the curvature endomorphism does not depend explicitly on the base-point $x$, therefore we omitted its indication.

A matrix-representation of the curvature is the curvature tensor $\kappa^{R(u, v)} = uv^T - vu^T$, 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 tensor $\kappa^{R(u, v)}$ coincides to the matrix representation of the 3-vector $u \wedge v$, namely $\kappa^{R(u, v)} = u \wedge v$.

The exponential map is expressed as
\[
\exp_y(w) := \begin{cases} 
  y \cos(\|w\|) + w \frac{\operatorname{sinc}(\|w\|)}{\operatorname{sinc}(d(x, y))} & \text{if } w \neq 0, \\
  y & \text{if } w = 0,
\end{cases}
\] (87)
therefore, its inverse, the logarithmic map, is expressed by (adapted from [11]):
\[
\log_y x = \frac{(I_n - yy^T)x}{\operatorname{sinc}(d(x, y))},
\] (88)
where it is assumed that the points $x, y \in \mathbb{S}^{n-1}$ are not antipodal. In the above formulas, the symbol $\operatorname{sinc} : \mathbb{R} \to \mathbb{R}$ denotes the cardinal sine function and the geodesic distance between two points $x, y \in \mathbb{S}^{n-1}$ associated to the canonical metrics reads
\[
d(x, y) = \| \log_y x \|_y = \cos(x^T y),
\] (89)
where the inverse cosine function is supposed to range in $[0, \pi]$.

Matrix representations of the derivatives of the exponential and the logarithmic maps are computed to be
\[
\begin{align*}
\left[\!(d \exp)_y,w\!\right] &= \cos(\|w\|)I_n - \operatorname{sinc}(\|w\|)yw^T, \\
\left[\!(d_x \log)_y,x\!\right] &= \frac{I_n - yy^T}{\operatorname{sinc}^2 d(x, y)} \left[\!\operatorname{sinc}(d(x, y))I_n - \left(\operatorname{sinc}' d(x, y)\right) \cos'(x^T y)xy^T\!\right],
\end{align*}
\] (90
and 91)
where the prime denotes derivation. The corresponding matrix representation of the endomorphism $\mathcal{R}^0$, upon defining $D_{x,y,\theta} := \left[\!(d \exp)_y,\theta \log_y x\!\right]$, may be written as:
\[
\left[\mathcal{R}^0\right] = D_{x,y,\theta} \left\{ (\log_y x)x^T \left[\!(d_x \log)_y,x\!\right]^T - \left[\!(d_x \log)_y,x\!\right] \log_y^T x \right\} D_{x,y,\theta}^T,
\] (92)
which is clearly a skew-symmetric matrix.

Compliance with ethical standards. Conflict of interest. The author declare that he has no conflict of interest.

Acknowledgment. Part of this research work was carried out while the author was on a temporary leave at the National Center for Theoretical Sciences (NCTS) at the Department of Mathematics of the National Taiwan University (NTU) in Taipei in August 2016, under a “Research in Pairs” cooperation project. This research work was then continued in 2017 during a visit to the Tokyo University of Agriculture and Technology.
REFERENCES

[1] W. Ambrose and I. M. Singer, A theorem on holonomy, Transactions of the American Mathematical Society, 75 (1953), 428–443.

[2] R. Anirudh, V. Venkataraman and P. Turaga, A generalized Lyapunov feature for dynamical systems on Riemannian manifolds, Proceedings of the 1st International Workshop on Differential Geometry in Computer Vision for Analysis of Shapes, Images and Trajectories, (2015), 4.1–4.10.

[3] A. Bejenaru and C. Udrisete, Multivariate optimal control with payoffs defined by submanifold integrals, Symmetry, 11 (2019), 893.

[4] F. Bullo, R. M. Murray and A. Sarti, Control on the sphere and reduced attitude stabilization, IFAC Proceedings Volumes, 28 (1995), 495–501.

[5] J. Casey, Parallel transport of a vector on a surface, Exploring Curvature, (1996), 250–263.

[6] Y. Chitour and M. Sigalotti, On the controllability of the Dubins’ problem for surfaces, Proceedings of the IFAC System, Structure and Control Conference, (2004), 563–565.

[7] Y. Chitour and M. Sigalotti, On the controllability of the Dubins problem for surfaces. I Nonnegative curvature, Journal of Geometric Analysis, 15 (2005), 565–587.

[8] Y. Chitour and M. Sigalotti, On the controllability of the Dubins problem for surfaces. II Negative curvature, SIAM Journal on Control and Optimization, 45 (2006), 457–482.

[9] R. Dandoloff, Berry’s phase and Fermi-Walker parallel transport, Physics Letters A, 139 (1989), 19–20.

[10] S. Eftekhar Azam, S. Mariani and N. K. A. Attari, Online damage detection via a synergy of proper orthogonal decomposition and recursive Bayesian filters, Nonlinear Dynamics, 89 (2017), 1489–1511.

[11] S. Fiori, On vector averaging over the unit hyphersphere, Digital Signal Processing, 19 (2009), 715–725.

[12] S. Fiori, Nonlinear damped oscillators on Riemannian manifolds: Fundamentals, Journal of Systems Science and Complexity, 29 (2016), 22–40.

[13] S. Fiori, Non-delayed synchronization of non-autonomous dynamical systems on Riemannian manifolds and its applications, Nonlinear Dynamics, 94 (2018), 3077–3100.

[14] S. Fiori, Nonlinear damped oscillators on Riemannian manifolds: Numerical simulation, Communications in Nonlinear Science and Numerical Simulation, 47 (2017), 207–222.

[15] R. Fuentes, G. Hicks and J. Osborne, The spring paradigm in tracking control of simple mechanical systems, Automatica, 47 (2011), 993–1000.

[16] F. A. Goodarzi, D. Lee and T. Lee, Geometric control of a quadrotor UAV transporting a payload connected via flexible cable, International Journal of Control, Automation, and Systems, 13 (2015), 1486–1498.

[17] M.-D. Hua, T. Hamel, R. Mahony and J. Trumpf, Gradient-like observer design on the special Euclidean group SE(3) with system outputs on the real projective space, Proceedings of the 54th IEEE Conference on Decision and Control, (2015), 2139–2145.

[18] M. Kabiri, H. Atrianfar and M. B. Menhaj, Trajectory tracking of a class of under-actuated thrust-propelled vehicle with uncertainties and unknown disturbances, Nonlinear Dynamics, 90 (2017), 1695–1706.

[19] L. M. Lopes, S. Fernandes and C. Grácio, Complete synchronization and delayed synchronization in couplings, Nonlinear Dynamics, 79 (2015), 1615–1624.

[20] L. Lü, C. Li, S. Liu, Z. Wang, J. Tian and J. Gu, The signal synchronization transmission among uncertain discrete networks with different nodes, Nonlinear Dynamics, 81 (2015), 801–809.

[21] C. W. Misner, K. S. Thorne and J. A. Wheeler, Gravitation, W.H. Freeman, 1973.

[22] A. Mortada, P. Kokkonen and Y. Chitour, Rolling manifolds of different dimensions, Acta Applicandae Mathematicae, 139 (2015), 105–131.

[23] K. Ojo, S. Ogunjo and A. Olagundoye, Projective synchronization via active control of identical chaotic oscillators with parametric and external excitation, International Journal of Nonlinear Science, 24 (2017), 76–83.

[24] J. M. Osborne and G. P. Hicks, The geodesic spring on the Euclidean sphere with parallel-transport-based damping, Notices of the AMS, 60 (2013), 544–556.

[25] P. Piras, L. Teresi, L. Travescetti, V. Varano, S. Gabriele, T. Kotsakis, P. Raia, P. E. Puddu and M. Scalici, The conceptual framework of ontogenetic trajectories: parallel transport allows the recognition and visualization of pure deformation patterns, Evolution & Development, 18 (2016), 182–200.
[26] W. H. Press, S. A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Third edition. Cambridge University Press, Cambridge, 2007.

[27] H. Reckziegel and E. Wilhelmus, How the curvature generates the holonomy of a connection in an arbitrary fibre bundle, *Results in Mathematics*, 49 (2006), 339–359.

[28] T. Soyfıdan, H. Parlatici and M. A. Gungör, On the quaternionic curve according to parallel transport frame, *TWMS Journal of Pure and Applied Mathematics*, 4 (2013), 194–203.

[29] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc., Wilmington, Del., 1979.

[30] Z. G. Ying and W. Q. Zhu, Optimal bounded control for nonlinear stochastic smart structure systems based on extended Kalman filter, *Nonlinear Dynamics*, 90 (2017), 105–114.

[31] W. Yu and Z. Pan, Dynamical equations of multibody systems on Lie groups, *Advances in Mechanical Engineering*, 7 (2015), 1–9.

[32] Z. Zhang, Z. Ling and A. Sarlette, Modified integral control globally counters symmetry-breaking biases, *Symmetry*, 11 (2019), 639.

[33] Y. Zou, Adaptive trajectory tracking control approach for a model-scaled helicopter, *Nonlinear Dynamics*, 83 (2016), 2171–2181.

Received August 2019; revised March 2020.

E-mail address: s.fiori@staff.univpm.it