High Order Smoothing Splines *versus* Least Squares Problems on Riemannian Manifolds

L. Machado F. Silva Leite K. Krakowski

May 9, 2008

Abstract

The aim of this work is to establish the generalization of the classical least squares problem on Euclidean spaces introduced by Lagrange to more general Riemannian manifolds. We start with the formulation of a high order variational problem on a Riemannian manifold, depending on a smoothing parameter, that generate smoothing geometric splines fitting a given data. This formulation is related to the definition of geometric polynomials on manifolds adopted about two decades ago.

The Riemannian mean of the given points is then achieved as a limiting process and moreover, when we particularize the Riemannian manifold to be an Euclidean space, the polynomial curve that is the solution of the classical least squares problem also arises as a limiting process.

These astounding facts support our strong belief that this is the best generalization of the classical least squares problem to Riemannian manifolds.

1 Introduction

Curve fitting techniques on Euclidean spaces are well known in the literature, being the classical least squares problems the most common. However, these fitting techniques reveal to be insufficient since the most part of the mechanical systems arising in modern applications have components that are manifolds such as Lie groups or symmetric spaces. Such is the case of the trajectory planning problem arising in robotics, aeronautics and air traffic control.

2 Preliminaries

In what follows, M denotes an $n-$dimensional Riemannian manifold endowed with its Riemannian connection (the Levi-Civita connection), that we denote by ∇. Given $p \in M$, T_pM denotes, as usual, the tangent space of M at p and we use the notation $\langle \cdot, \cdot \rangle$ to represent the inner product in T_pM. TM stands for the tangent bundle of M and is therefore the disjoint union $\bigcup_{p\in M} T_pM$. By a vector field in M we meant a mapping $X : M \to TM$ that assigns to each $p \in M$ a vector $X_p \in T_pM$.

A curve c in M is simply a parameterized curve $c : I \subset \mathbb{R} \to M$ from an interval I of real numbers to M. A vector field V along a curve c is, therefore, a mapping that assigns to each $t \in I$, the vector $V(t) \in T_{c(t)}M$. The velocity vector field of c, that we denote by $\dfrac{dc}{dt}$, is such an example. If V is induced by some vector field X in M, that is, if $V(t) = X_{c(t)}$, then we define the covariant derivative of V along c as being

$$\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} X. \tag{2.1}$$

More generally, we have

$$\frac{D^m V}{dt^m} = \frac{D^{m-1}}{dt^{m-1}}\Big(\frac{DV}{dt}\Big), \;\; \forall\, m \geq 2.$$

The Riemannian connection ∇ is the unique affine connection that is compatible with the Riemannian metric and therefore, if V and W are smooth vector fields along a curve c, then

$$\frac{d}{dt}\Big\langle V, W\Big\rangle = \Big\langle \frac{DV}{dt}, W\Big\rangle + \Big\langle V, \frac{DW}{dt}\Big\rangle. \tag{2.2}$$

The previous equality can be seen as a particular case of the more general property that is stated in the following lemma.

Lemma 2.1.

$$\Big\langle \frac{D^p V}{dt^p}, \frac{D^q W}{dt^q}\Big\rangle = \sum_{l=1}^{p} (-1)^{l-1} \frac{d}{dt}\Big\langle \frac{D^{p-l} V}{dt^{p-l}}, \frac{D^{q+l-1} W}{dt^{q+l-1}}\Big\rangle + (-1)^p \Big\langle V, \frac{D^{p+q} W}{dt^{p+q}}\Big\rangle,$$

where $p, q \in \mathbb{N}_0$.

A vector field V is said to be a parallel vector field along a curve c if

$$\frac{DV}{dt} = 0. \tag{2.3}$$

Taking into account the existence and uniqueness theorem for ordinary differential equations, it can be easily seen that given $V_0 \in T_{c(0)}M$, there exists a unique parallel vector field V along c such that $V(0) = V_0$. This vector field is called the parallel translate of V_0 along c. Thus, we can establish a linear isomorphism between tangent spaces, called the parallel transport,

$$\begin{array}{rccl} P_{0,t}: & T_{c(0)}M & \longrightarrow & T_{c(t)}M \\ & V_0 & \longmapsto & P_{0,t}(V_0) = V(t) \end{array}.$$

being $V(t)$ the unique parallel translate of V_0 along c.

By definition, a geodesic c is a smooth curve whose velocity vector is a parallel vector field along c. That is,

$$\frac{D}{dt}\Big(\frac{dc}{dt}\Big) = 0.$$

The above condition can also be written for simplicity as $\dfrac{D^2 c}{dt^2} = 0$. Therefore, geodesics are constant speed curves and they can be parameterized explicitly by

$$c(t) = \exp_p(tv),$$

where $v \in T_pM$ and $\exp : TM \to M$ stands for the exponential map in M, [19].

Although, in general, the exponential map is only a terminology, there are some special Riemannian manifolds where it can be explicitly defined. In Euclidean spaces, geodesics for the usual Riemannian metric, are the straight lines and therefore, the exponential map is given by

$$\exp_p(tv) = p + tv.$$

On the other hand, geodesics in the unit $n-$sphere S^n, are the great arc circles for the Riemannian metric induced by the usual inner product in $\mathbb{R}^{n+1}$, and therefore,

$$\exp_p(tv) = p\cos\big(t\|v\|\big) + \frac{v}{\|v\|}\sin\big(t\|v\|\big). \tag{2.4}$$

For the case of connected and compact Lie groups, geodesics can be defined as

$$\exp_p(tv) = p\,V\mathrm{e}^{tv}, \tag{2.5}$$

where e^{tv} denotes the sum of the power series $\mathrm{e}^{tv} = \sum_{m=0}^{+\infty} \frac{t^m v^m}{m!}$.

Since $v = \frac{dc}{dt}(0) \in T_{c(0)}M$, using the definition of the parallel transport given above, it follows that

$$P_{0,t}(v) = \frac{dc}{dt}(t).$$

When M is geodesically complete, any two points p and q, sufficiently close, can be joined by a unique minimal geodesic arc that can be parameterized explicitly as in [13], by

$$c(s) = \exp_p\big(s\exp_p^{-1}(q)\big), \quad s \in \big[0,1\big]. \tag{2.6}$$

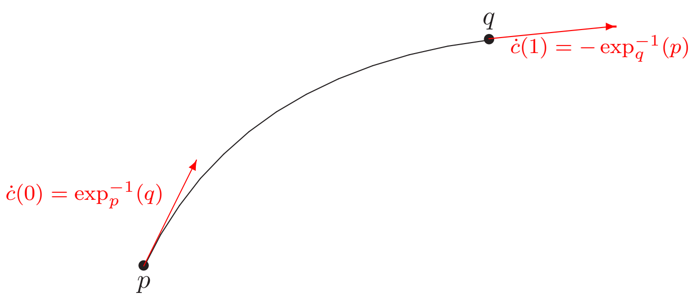

Figure 1: The minimal geodesic joining points p to q.

In this case, the distance between points p and q, is therefore

$$d(p,q) = \Big\langle \exp_p^{-1}(q), \exp_p^{-1}(q)\Big\rangle^{\frac{1}{2}},$$

and M becomes a complete metric space when endowed with the metric induced by the above distance function.

A subset $C \subset M$ is said to be geodesically convex if all minimizing geodesics that start and end in C must lie entirely in C, [8, 5].

Keeping the same terminology of do Carmo [8], we adopt the following definition for the curvature tensor in M:

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

where X, Y and Z are smooth vector fields in M.

Let $\alpha : (x,y) \in \mathbb{R}^2 \mapsto \alpha(x,y) \in M$ be a smooth parameterized surface in M. In spite of the following symmetry condition (Lee [16])

$$\frac{D}{\partial x}\Big(\frac{\partial \alpha}{\partial y}\Big) = \frac{D}{\partial y}\Big(\frac{\partial \alpha}{\partial x}\Big), \tag{2.7}$$

the two covariant differentiation operators $\dfrac{D}{\partial y}$ and $\dfrac{D}{\partial y}$ do not commute in general. The extent of noncommutativity of these two operators is given by the curvature tensor as it is shown in the next lemma.

Lemma 2.2. [19] *If V is a vector field along the parameterized surface α, then*

$$\frac{D}{\partial y}\frac{D}{\partial x}V = \frac{D}{\partial x}\frac{D}{\partial y}V + R\Big(\frac{\partial \alpha}{\partial y}, \frac{\partial \alpha}{\partial x}\Big)V. \tag{2.8}$$

Using high order covariant differentiation, it can also be established the more general result, that can be found in [3].

Proposition 2.3.

$$\frac{D}{\partial y}\Big(\frac{D^m V}{\partial x^m}\Big) = \frac{D^m}{\partial x^m}\Big(\frac{DV}{\partial y}\Big) + \sum_{j=2}^{m} \frac{D^{m-j}}{\partial t^{m-j}} R\Big(\frac{\partial \alpha}{\partial y}, \frac{\partial \alpha}{\partial x}\Big)\frac{D^{j-1}\alpha}{\partial x^{j-1}}. \tag{2.9}$$

The curvature tensor satisfies several symmetry relations that will be used throughout the paper and are listed below.

Lemma 2.4. [19] *If X, Y, Z and W are smooth vector fields, the curvature tensor R satisfies the following symmetry relations:*

1. $R(X,Y)Z = -R(Y,X)Z$;
2. $R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0$;
3. $\big\langle R(X,Y)Z, W\big\rangle = -\big\langle R(X,Y)W, Z\big\rangle$;
4. $\big\langle R(X,Y)Z, W\big\rangle = \big\langle R(W,Z)Y, X\big\rangle$.

Finally, given a point $p \in M$, and a two dimensional subspace Ξ of T_pM, if $\big\{X, Y\big\}$ is any basis of Ξ, the real number

$$\Delta(\Xi) = \frac{\Big\langle R(X,Y)Y, X\Big\rangle}{\sqrt{\|X\|^2\|Y\|^2 - \langle X,Y\rangle^2}},$$

denotes the sectional curvature of Ξ at p.

2.1 Riemannian Mean

In Euclidean spaces there are several concepts of means [20], each of them with numerous applications in different areas. Nevertheless, the most common is indeed the arithmetic mean, also know as the center of mass, centroid or barycenter. For the set of points $p_0, \ldots, p_N$, belonging to the Euclidean space $\mathbb{R}^n$, it is simply defined as

$$\bar{p} = \frac{1}{N+1} \sum_{i=0}^{N} p_i, \tag{2.10}$$

due to the straight relationship between points and vectors on Euclidean spaces.

The above formula has not a straightforward generalization to more general Riemannian manifolds. However, the arithmetic mean (2.10) in Euclidean spaces can also be interpreted as the point p that minimizes the sum of the squared Euclidean distances between p and each p_i, that is, is the unique solution of the following minimization problem:

$$\min_{p \in \mathbb{R}^n} \sum_{i=0}^{N} \left\| p - p_i \right\|^2.$$

Now, it is already possible to generalize the above formulation can to a more general Riemannian manifold M, if we simply replace the Euclidean distance by the geodesic distance. Then, we can define the Riemannian mean of the points $p_0, \ldots, p_N$ lying in M, as being the set of points $p \in M$ that yield the minimum value for the function

$$\Phi(p) = \sum_{i=0}^{N} d^2(p, p_i). \tag{2.11}$$

It has been already proved in the literature ([13] and [14]), that a necessary condition for $p \in M$ to be a local minimum for Φ is that

$$\sum_{i=0}^{N} \exp_p^{-1}(p_i) = 0. \tag{2.12}$$

Contrary to what happens on Euclidean spaces, we have no guarantee that the Riemannian mean of a set of points is unique. If we think on two antipodal points on the sphere S^2 (figure 2), it is easy to check that all the points lying in the equator yield the minimum value for the function (2.11).

However, when the points are sufficiently close enough, it has been proved in Karcher [13], that the Riemannian mean of the given points is unique.

Theorem 2.5. [13] *If B_ρ is a convex geodesic ball in M, with radius $\rho < \frac{\pi}{4}\Delta^{-\frac{1}{2}}$, being $\Delta > 0$ the maximum value of the sectional curvature in B_ρ, then function Φ is convex in B_ρ and it has a unique point of local minimum in B_ρ.*

The above result has been already extended by several authors for some particular symmetric spaces, like for instance, the Lie group of rotations [14], [12], or the unit $n-$sphere [1].

In this paper, we will present an alternative way to obtain the Riemannian mean of a given set of points in M based in the formulation of a variational problem whose solutions are broken geodesics fitting those points.

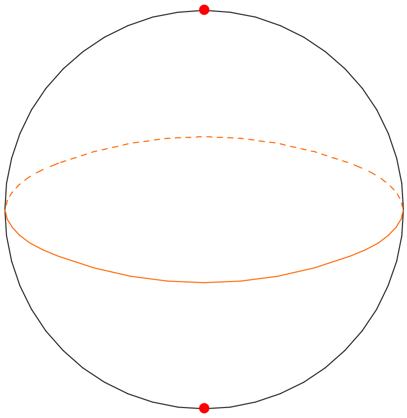

Figure 2: The mean of two antipodal points in S^2.

2.2 High Order Polynomials on Riemannian Manifolds

Polynomials on Euclidean spaces are well known behaviored curves that have a wide range of applications. Actually, interpolating splines based on cubic polynomials are the most used in approximation theory and the classical least squares problems introduced by Lagrange (1736-1813), also based in Euclidean polynomials, are a typical tool in the context of fitting curves.

Since our aim here is to establish the generalization of the classical least squares problems to more general Riemannian manifolds, the first step is to define polynomials on manifolds.

About two decades ago, cubic polynomials on Riemannian manifolds have been introduced by Noakes, Heinzinger and Paden [21], has being the extremal curves for the functional

$$L_2(\gamma) = \frac{1}{2}\int_0^T \Big\langle \frac{D^2\gamma}{dt^2}, \frac{D^2\gamma}{dt^2} \Big\rangle \, dt,$$

over an appropriate family of smooth curves $\gamma : [0,T] \to M$, satisfying some prescribed boundary conditions.

Analogously to what happens in Euclidean spaces [9], cubic polynomials on Riemannian manifolds also minimize changes in the acceleration , but only that component that is tangent to the manifold.

Later on, high order polynomials on Riemannian manifolds, also known as geometric polynomials, have been introduced in the literature by Camarinha et al. [3], as a generalization of the above and have been defined as the extremals for the functional

$$L_m(\gamma) = \frac{1}{2}\int_0^T \Big\langle \frac{D^m\gamma}{dt^m}, \frac{D^m\gamma}{dt^m} \Big\rangle \, dt. \tag{2.13}$$

Therefore, we say that a smooth curve $\gamma : [0,T] \to M$ is a geometric polynomial of degree $2m-1$, if it satisfies the following differential equation

$$\frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m} (-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt} = 0. \tag{2.14}$$

Cubic polynomials are therefore obtained by considering $m = 2$ in (2.14). Even for this particular case, the differential equation that becomes

$$\frac{D^4\gamma}{dt^4} + R\Big(\frac{D^2\gamma}{dt^2}, \frac{d\gamma}{dt}\Big)\frac{d\gamma}{dt} = 0,$$

is highly non-linear and in spite of the effort that has been taken by several researchers from different perspectives, many questions concerning the geometry and ways to compute them remain open. For more details, we mention [21], [2], [6], [7], [24], [4], [10], [22], and the references therein.

In the next lemma, we define an invariant along a geometric polynomial that will be useful to prove some of the results appearing in the next section that was derived independently in [17] and in [22].

Lemma 2.6. *The quantity*

$$I = \sum_{j=1}^{m-1}(-1)^{j-1}\Big\langle \frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^j\gamma}{dt^j}\Big\rangle + \frac{(-1)^{m-1}}{2}\Big\langle \frac{D^m\gamma}{dt^m}, \frac{D^m\gamma}{dt^m}\Big\rangle, \tag{2.15}$$

is preserved along a smooth curve satisfying (2.14)*.*

For the particular case when $m = 2$, the invariant (2.15) reduces to the invariant along a cubic polynomial derived in Camarinha et. al. [4].

Our next result provides an alternative characterization of the above invariant for the particular case when it vanishes identically, that will be useful to prove some of the results appearing in the next section.

Lemma 2.7. *If the invariant* (2.15) *vanishes identically along the geometric polynomial* (2.14)*, then*

$$\sum_{j=1}^{m-1}(-1)^{j-1} j \frac{d}{dt}\Big\langle \frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big\rangle = (-1)^m\Big(m - \frac{1}{2}\Big)\Big\langle \frac{D^m\gamma}{dt^m}, \frac{D^m\gamma}{dt^m}\Big\rangle. \tag{2.16}$$

3 Problem's Statement

Let Ω denote the set of all C^{m-1} paths $\gamma : [0,1] \to M$ such that $\gamma\Big|_{[t_i,t_{i+1}]}$ is smooth (C^∞) and therefore both the limits $\displaystyle\lim_{x\to t_i^+} \frac{D^k\gamma}{dt^k}(t)$ and $\displaystyle\lim_{x\to t_{i+1}^-} \frac{D^k\gamma}{dt^k}(t)$ are bounded, for all $k \in \mathbb{N}$.

We define the tangent space of Ω at a path γ, $T_\gamma\Omega$, as being the set of all C^{m-1} vector fields $W : [0,1] \to TM$ such that $W\Big|_{[t_i,t_{i+1}]}$ is smooth.

Given a set of points in M, $p_0, p_1, \ldots, p_N$, and a set of instants of time $0 = t_0 < t_1 < \cdots < t_N = 1$, let us consider the following variational problem

$$(\mathcal{P}) \qquad \min_{\gamma\in\Omega} J(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2\big(p_i, \gamma(t_i)\big) + \frac{\lambda}{2}\int_0^1 \Big\langle \frac{D^m\gamma}{dt^m}, \frac{D^m\gamma}{dt^m}\Big\rangle\, dt,$$

where λ denotes a positive real number that will play the role of a smoothing parameter as we will see sooner.

In order to simplify our exposition, let us denote by E the functional

$$E(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2\big(p_i, \gamma(t_i)\big),$$

Therefore, taking into account the definition of the functional L_m given in (2.13), we can write

$$J(\gamma) = E(\gamma) + \lambda L_m(\gamma).$$

In order to find the the first order necessary optimality conditions for problem $(\mathcal{P})$, we need to compute the first variation of J, that is,

$$\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big), \tag{3.17}$$

where $\alpha : \big]-\varepsilon,\varepsilon\big[\times \big[0,1\big] \longmapsto \alpha(u,t) \in M$ is any variation of γ.

Therefore, γ is an extremal for the functional J if and only its first variation (3.17) vanishes for all variations α of γ.

Let us pick up a variation α defined by

$$\alpha(u,t) = \exp_{\gamma(t)}\big(uW(t)\big), \tag{3.18}$$

for some variation vector field $W : \big[0,1\big] \to TM$ along γ lying in $T_\gamma\Omega$. Therefore,

$$W(t) = \frac{\partial \alpha}{\partial u}(0,t).$$

Since,

$$\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big) = \frac{\partial}{\partial u}\Big|_{u=0} E\big(\alpha(u,t)\big) + \lambda \frac{\partial}{\partial u}\Big|_{u=0} L\big(\alpha(u,t)\big),$$

we will start with the computation of $\displaystyle\frac{\partial}{\partial u}\Big|_{u=0} E\big(\alpha(u,t)\big)$.

For each $i = 0,\dots,N$, let us denote by

$$c_i(s) = \exp_{p_i}\big(s \exp_{p_i}^{-1}\big(\gamma(t_i)\big)\big), \tag{3.19}$$

the minimal geodesic joining the point p_i (at $s=0$) to the point $\gamma(t_i)$ (at $s=1$).

Introducing in (3.19) the variation α defined by (3.18), we obtain the parameterized surface in M, $c_i : \big[0,1\big] \times \big]-\varepsilon,\varepsilon\big[\longrightarrow M$, given by

$$c_i(s,u) = \exp_{p_i}\big(s \exp_{p_i}^{-1}\big(\alpha(u,t_i)\big)\big).$$

Therefore, we can define two family of curves

$$s \longmapsto c_i(s,u),$$

by setting u constant, and

$$u \longmapsto c_i(s,u),$$

by setting s constant, and consequently two family of vector fields

$$S_i(s,u) = \frac{\partial c_i}{\partial s}(s,u),$$

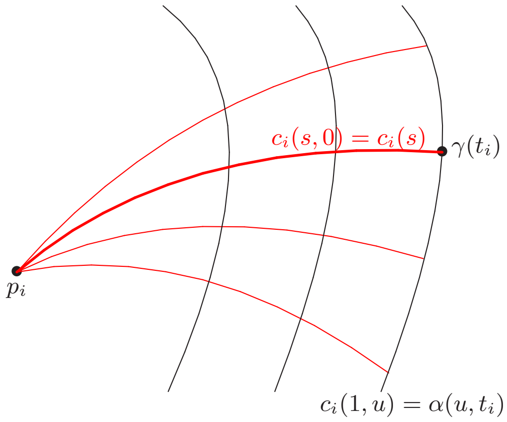

Figure 3: The parameterized surface c_i.

and, analogously,

$$U_i(s,u) = \frac{\partial c_i}{\partial u}(s,u).$$

Since for each fixed u, $s \longmapsto c_i(s,u)$ is a geodesic, S_i is a parallel vector field along that geodesic, which means that

$$\frac{DS_i}{\partial s}(s,u) = 0.$$

On the other hand, there exists a unique minimizing geodesic joining p_i to $\alpha(u,t_i)$ (see figure 3), so we can write

$$d^2\big(p_i, \alpha(u,t_i)\big) = \Big\langle S_i(s,u), S_i(s,u)\Big\rangle = \int_0^1 \Big\langle S_i(s,u), S_i(s,u)\Big\rangle\, ds.$$

Now, using the symmetry condition (s2.7) together with the compatibility condition (2.2), we can still write

$$\begin{aligned}
\frac{\partial}{\partial u}E\big(\alpha(u,t)\big) &= \sum_{i=0}^{N}\int_0^1 \Big\langle \frac{DS_i}{\partial u}(s,u), S_i(s,u)\Big\rangle\, ds \\
&= \sum_{i=0}^{N}\int_0^1 \Big\langle \frac{DU_i}{\partial s}(s,u), S_i(s,u)\Big\rangle\, ds \\
&= \sum_{i=0}^{N}\int_0^1 \frac{\partial}{\partial s}\Big\langle U_i(s,u), S_i(s,u)\Big\rangle\, ds \\
&= \sum_{i=0}^{N}\Big\langle U_i(1,u), S_i(1,u)\Big\rangle - \Big\langle U_i(0,u), S_i(0,u)\Big\rangle.
\end{aligned}$$

By setting $u = 0$, and taking into account that $S_i(1,0) = -\exp^{-1}_{\gamma(t_i)}(p_i)$ (figure 1), we get

$$\frac{\partial}{\partial u}\Big|_{u=0} E\big(\alpha(u,t)\big) = -\sum_{i=0}^{N}\Big\langle W(t_i), \exp^{-1}_{\gamma(t_i)}(p_i)\Big\rangle. \tag{3.20}$$

It remains the computation of the first variation of functional L. In this case, we will use lemma 2.1 and proposition 2.3 and follow the analogous steps as in [3] adapted to the current situation.

$$
\begin{aligned}
&\frac{\partial}{\partial u} L\big(\alpha(u,t)\big) = \\
&= \int_0^1 \Big\langle \frac{D}{\partial u}\Big(\frac{D^m \alpha}{\partial t^m}\Big), \frac{D^m \alpha}{\partial t^m} \Big\rangle \, dt \\
&= \int_0^1 \Big\langle \frac{D^m}{\partial t^m}\Big(\frac{\partial \alpha}{\partial u}\Big), \frac{D^m \alpha}{\partial t^m} \Big\rangle \, dt + \sum_{j=2}^{m} \int_0^1 \Big\langle \frac{D^{m-j}}{\partial t^{m-j}} R\Big(\frac{\partial \alpha}{\partial u}, \frac{\partial \alpha}{\partial t}\Big) \frac{D^{j-1}\alpha}{\partial t^{j-1}}, \frac{D^m \alpha}{\partial t^m} \Big\rangle \, dt \\
&= \sum_{l=1}^{m} (-1)^{l-1} \int_0^1 \frac{\partial}{\partial t} \Big\langle \frac{D^{m-l}}{\partial t^{m-l}}\Big(\frac{\partial \alpha}{\partial u}\Big), \frac{D^{m+l-1}\alpha}{\partial t^{m+l-1}} \Big\rangle \, dt + (-1)^m \int_0^1 \Big\langle \frac{\partial \alpha}{\partial u}, \frac{D^{2m}\alpha}{\partial t^{2m}} \Big\rangle \, dt \\
&\quad + \sum_{j=2}^{m-1} \sum_{l=1}^{m-j} (-1)^{l-1} \int_0^1 \frac{\partial}{\partial t} \Big\langle \frac{D^{m-j-l}}{\partial t^{m-j-l}} R\Big(\frac{\partial \alpha}{\partial u}, \frac{\partial \alpha}{\partial t}\Big) \frac{D^{j-1}\alpha}{\partial t^{j-1}}, \frac{D^{m+l-1}\alpha}{\partial t^{m+l-1}} \Big\rangle \, dt \\
&\quad + \sum_{j=2}^{m} (-1)^{m-j} \int_0^1 \Big\langle R\Big(\frac{\partial \alpha}{\partial u}, \frac{\partial \alpha}{\partial t}\Big) \frac{D^{j-1}\alpha}{\partial t^{j-1}}, \frac{D^{2m-j}\alpha}{dt^{2m-j}} \Big\rangle \, dt.
\end{aligned}
$$

By letting $u = 0$ in the above expression and taking into account property 4 of the curvature tensor listed in lemma 2.4, we get

$$
\begin{aligned}
&\frac{\partial}{\partial u}\Big|_{u=0} L\big(\alpha(u,t)\big) = \\
&= \sum_{l=1}^{m} \sum_{i=0}^{N-1} (-1)^{l-1} \Big\langle \frac{D^{m-l}W}{dt^{m-l}}, \frac{D^{m+l-1}\gamma}{dt^{m+l-1}} \Big\rangle \Big|_{t_i^+}^{t_{i+1}^-} \\
&\quad + \sum_{j=2}^{m-1} \sum_{l=1}^{m-j} \sum_{i=0}^{N-1} (-1)^{l-1} \Big\langle \frac{D^{m-j-l}}{dt^{m-j-l}} R\Big(W, \frac{d\gamma}{dt}\Big) \frac{D^{j-1}\gamma}{dt^{j-1}}, \frac{D^{m+l-1}\gamma}{dt^{m+l-1}} \Big\rangle \Big|_{t_i^+}^{t_{i+1}^-} \\
&\quad + (-1)^m \int_0^1 \Big\langle \frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m} (-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big) \frac{d\gamma}{dt}, W \Big\rangle \, dt
\end{aligned}
\tag{3.21}
$$

Putting together (3.20) and (3.21), we obtain the desired first variation of

J,

$$
\begin{aligned}
&\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big) = \\
&= \sum_{l=1}^{m-1}\sum_{i=0}^{N}(-1)^l\lambda\Big\langle \frac{D^{m-l}W}{dt^{m-l}}(t_i), \frac{D^{m+l-1}\gamma}{dt^{m+l-1}}(t_i^+) - \frac{D^{m+l-1}\gamma}{dt^{m+l-1}}(t_i^-)\Big\rangle \\
&\quad + \sum_{i=0}^{N}\Big\langle W(t_i), (-1)^m\lambda\Big[\frac{D^{2m-1}\gamma}{dt^{2m-1}}(t_i^+) - \frac{D^{2m-1}\gamma}{dt^{2m-1}}(t_i^-)\Big] - \exp^{-1}_{\gamma(t_i)}(p_i)\Big\rangle \\
&\quad + \sum_{j=2}^{m-1}\sum_{l=1}^{m-j}\sum_{i=0}^{N-1}(-1)^{l-1}\lambda\Big\langle \frac{D^{m-j-l}}{dt^{m-j-l}}R\Big(W,\frac{d\gamma}{dt}\Big)\frac{D^{j-1}\gamma}{dt^{j-1}}, \frac{D^{m+l-1}\gamma}{dt^{m+l-1}}\Big\rangle\Big|_{t_i^+}^{t_{i+1}^-} \\
&\quad + (-1)^m\int_0^1 \lambda\Big\langle \frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m}(-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt}, W\Big\rangle\, dt,
\end{aligned}
\tag{3.22}
$$

and we can therefore establish one of our main results.

Theorem 3.1. *A necessary condition for $\gamma \in \Omega$ to be a solution for problem $(\mathcal{P})$ is that γ is of class C^{2m-2} in the whole interval $\big[0,1\big]$, satisfies for each $i = 0,\dots N-1$, and $t \in \big[t_i, t_{i+1}\big]$, the differential equation*

$$
\frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m}(-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt} = 0, \tag{3.23}
$$

and at the knot points t_i, for $i = 0,\dots,N$, it also satisfies the following conditions

$$
\left\{
\begin{array}{ll}
\dfrac{D^j\gamma}{dt^j}(t_i^+) - \dfrac{D^j\gamma}{dt^j}(t_i^-) = 0, & m \le j \le 2m-2 \\[2ex]
\dfrac{D^{2m-1}\gamma}{dt^{2m-1}}(t_i^+) - \dfrac{D^{2m-1}\gamma}{dt^{2m-1}}(t_i^-) = \dfrac{(-1)^m}{\lambda}\exp^{-1}_{\gamma(t_i)}(p_i) &
\end{array}
\right. , \tag{3.24}
$$

where we assume for shorten of notation that $\dfrac{D^j\gamma}{dt^j}(t_0^-) = \dfrac{D^j\gamma}{dt^j}(t_N^+) = 0$, for $j = m,\dots,2m-1$.

Proof. In order for $\gamma \in \Omega$ to be an extremal for functional J, its first variation has to vanish for all variations α given by (3.18).

Let us consider a variation vector field W defined as

$$
W(t) = (-1)^m F(t)\Big[\frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m}(-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt}\Big],
$$

where $F : [0,1] \to \mathbb{R}$ is a positive piecewise smooth function satisfying, for each $i = 0,\dots,N$,

$$
F(t_i) = F'(t_i) = \cdots = F^{(m-1)}(t_i) = 0.
$$

For this choice of the variation vector field W, we get

$$
\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big) = \int_0^1 F(t)\Big\|\frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m}(-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt}, W\Big\|^2 dt,
$$

and this vanishes identically if and only if, for each $i = 0, \dots, N$ and $t \in \big[t_i, t_{i+1}\big]$,

$$\frac{D^{2m}\gamma}{dt^{2m}} + \sum_{j=2}^{m} (-1)^j R\Big(\frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big)\frac{d\gamma}{dt} = 0.$$

Now, if one considers a variation vector field W satisfying, for each $i = 0, \dots, N$,

$$W(t_i) = \frac{DW}{dt}(t_i) = \cdots = \frac{D^{m-2}W}{dt^{m-2}}(t_i) = 0,$$

and

$$\frac{D^{m-1}W}{dt^{m-1}}(t_i) = \frac{D^m\gamma}{dt^m}(t_i^-) - \frac{D^m\gamma}{dt^m}(t_i^+),$$

we get

$$\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big) = \sum_{i=0}^{N} \lambda \Big\| \frac{D^m\gamma}{dt^m}(t_i^+) - \frac{D^m\gamma}{dt^m}(t_i^-)\Big\|^2,$$

which vanishes if and only if γ is of class C^m in the whole interval $\big[0,1\big]$.

Now, if we choose a variation vector field W such that for each $i = 0, \dots, N$,

$$W(t_i) = \frac{DW}{dt}(t_i) = \cdots = \frac{D^{m-3}W}{dt^{m-3}}(t_i) = 0,$$

and

$$\frac{D^{m-2}W}{dt^{m-2}}(t_i) = \frac{D^{m+1}\gamma}{dt^{m+1}}(t_i^+) - \frac{D^{m+1}\gamma}{dt^{m+1}}(t_i^-),$$

it can easily be seen that if γ is a solution for problem $(\mathcal{P})$, then γ has to be of class C^{m+1} in the whole interval $\big[0,1\big]$.

Proceeding analogously, it can be proved that in order for γ to be an extremal for function J, γ has to be of class C^{2m-2} in the whole interval $\big[0,1\big]$.

If we finally choose the variation vector field W satisfying

$$W(t_i) = (-1)^m \lambda\Big[\frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^+\big) - \frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^-\big)\Big] - \exp^{-1}_{\gamma(t_i)}\big(p_i\big),$$

for each $i = 0, \dots, N$, we conclude that, for this case,

$$\frac{\partial}{\partial u}\Big|_{u=0} J\big(\alpha(u,t)\big) = \sum_{i=0}^{N} \Big\|(-1)^m \lambda\Big[\frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^+\big) - \frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^-\big)\Big] - \exp^{-1}_{\gamma(t_i)}\big(p_i\big)\Big\|^2,$$

and, therefore,

$$\frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^+\big) - \frac{D^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^-\big) = \frac{(-1)^m}{\lambda} \exp^{-1}_{\gamma(t_i)}\big(p_i\big),$$

which completes the proof. □

Remark 3.1. *From the above theorem, we can see that solutions for the variational problem* $(\mathcal{P})$ *are obtained by piecing together geometric polynomials of degree* $2m-1$ *in each subinterval* $\big[t_i, t_{i+1}\big]$*, that, according to the regularity conditions* (3.24)*, fit the given points* p_i *at the given times* t_i*. This is the reason why we call them smoothing geometric splines.*

Proposition 3.2. *If in conditions* (3.23)-(3.24) *of theorem* (3.1)*, we let the parameter* λ *going to*

(a) 0*, then the smoothing geometric splines approach an interpolating spline that passes through each point* p_i *at each time* t_i*;*

(b) $+\infty$*, then the smoothing geometric splines approach a smooth curve in the whole interval* $[0,1]$ *fitting the given points at the given times, and satisfying the differential equation*

$$\frac{D^m\gamma}{dt^m} = 0. \tag{3.25}$$

Proof. Property (a) follows immediately if one multiply both terms of the the last equation in (3.24), and then let λ going to 0.

To prove property (b), let us consider λ going to $+\infty$ in the last equation of (3.24). In that case, the curve γ becomes of class C^{2m-1} in the whole interval $[0,1]$, and satisfy

$$\frac{D^k\gamma}{dt^k}(0) = 0, \tag{3.26}$$

for $k = m, \ldots, 2m-1$.

According to the theory of ordinary differential equations, since γ satisfies the differential equation (3.23) of order $2m$ in each subinterval $[t_i, t_{i+1}]$ and is of class C^{2m-1} in the entire interval $[0,1]$, it has to satisfy (3.23), for all $t \in [0,1]$.

Now, taking into account the boundary conditions (3.26), it is easy to see that the invariant I along the geometric polynomial γ satisfying (3.23) vanishes identically. That is,

$$\sum_{j=1}^{m-1}(-1)^{j-1}\Big\langle \frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^j\gamma}{dt^j}\Big\rangle + \frac{(-1)^{m-1}}{2}\Big\langle \frac{D^m\gamma}{dt^m}, \frac{D^m\gamma}{dt^m}\Big\rangle = 0$$

Using lemma 2.7, we can conclude that the real function

$$t \longmapsto \sum_{j=1}^{m-1}(-1)^{j-1} j \Big\langle \frac{D^{2m-j}\gamma}{dt^{2m-j}}, \frac{D^{j-1}\gamma}{dt^{j-1}}\Big\rangle,$$

is a monotonous function in the interval $[0,1]$ (non-increasing for odd values of m and non-decreasing, otherwise).

In both cases, according to the boundary conditions (3.26), the above function vanishes identically in $[0,1]$, and therefore,

$$\frac{D^m\gamma}{dt^m}(t) = 0, \quad \forall t \in [0,1].$$

□

Remark 3.2. *Smooth curves satisfying the differential equation* (3.25) *can be obtained by rolling (without slip or twist) a manifold on its tangent space* $T_{\gamma(0)}M$ *along an Euclidean polynomial of degree* $m-1$*, as it was shown recently by Hüper and Silva Leite in* [11].

The previous results show that for the particular case when $m = 1$ and the smoothing parameter λ goes to infinity in the system of equations (3.23)-(3.24), the smoothing geometric splines approach a single point. What we will prove next is that this point turns out to be the Riemannian mean of the given points p_i, if we assume in advance that their Riemannian mean exists and is a singleton.

Theorem 3.3. *When $m = 1$ and λ goes to $+\infty$ in the conditions* (3.23)-(3.24), of theorem 3.1, then the smoothing geometric splines approach the Riemannian mean of the given points p_i.

Proof. For the particular case when $m = 1$, the differential equation (3.23) becomes

$$\frac{D^2\gamma}{dt^2} = 0, \tag{3.27}$$

and the regularity conditions (3.24) reduce simply to

$$\begin{aligned}
&\frac{d\gamma}{dt}(t_0^+) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_0)}(p_0)\\
&\frac{d\gamma}{dt}(t_1^+) - \frac{d\gamma}{dt}(t_1^-) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_1)}(p_1)\\
&\qquad\vdots\\
&\frac{d\gamma}{dt}(t_{N-1}^+) - \frac{d\gamma}{dt}(t_{N-1}^-) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_{N-1})}(p_{N-1})\\
&\frac{d\gamma}{dt}(t_N^-) = \frac{1}{\lambda}\exp^{-1}_{\gamma(t_N)}(p_N)
\end{aligned} \quad . \tag{3.28}$$

Since γ is a geodesic in each subinterval $\big[t_i, t_{i+1}\big]$, let us denote by P_i the parallel transport along γ in that subinterval. This means that,

$$\frac{d\gamma}{dt}(t_{i+1}^-) = P_i\Big(\frac{d\gamma}{dt}(t_i^+)\Big), \tag{3.29}$$

and the regularity conditions (3.28) may be written as

$$\begin{aligned}
&\frac{d\gamma}{dt}(t_0^+) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_0)}(p_0)\\
&\frac{d\gamma}{dt}(t_1^+) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_1)}(p_1) - \frac{1}{\lambda}P_0\Big(\exp^{-1}_{\gamma(t_0)}(p_0)\Big)\\
&\quad\vdots\\
&\frac{d\gamma}{dt}(t_{N-1}^+) = -\frac{1}{\lambda}\exp^{-1}_{\gamma(t_{N-1})}(p_{N-1}) - \frac{1}{\lambda}P_{N-2}\Big(\exp^{-1}_{\gamma(t_{N-2})}(p_{N-2})\Big) - \cdots\\
&\qquad\qquad -\frac{1}{\lambda}\Big(P_{N-2}\circ P_{N-3}\circ\cdots\circ P_0\Big)\Big(\exp^{-1}_{\gamma(t_0)}(p_0)\Big)\\
&-\frac{1}{\lambda}P_{N-1}\Big(\exp^{-1}_{\gamma(t_{N-1})}(p_{N-1})\Big) - \frac{1}{\lambda}\Big(P_{N-1}\circ P_{N-2}\Big)\Big(\exp^{-1}_{\gamma(t_{N-2})}(p_{N-2})\Big) - \cdots\\
&-\frac{1}{\lambda}\Big(P_{N-1}\circ P_{N-2}\circ\cdots\circ P_0\Big)\Big(\exp^{-1}_{\gamma(t_0)}(p_0)\Big) = \frac{1}{\lambda}\exp^{-1}_{\gamma(t_N)}(p_N)
\end{aligned} \quad . \tag{3.30}$$

When λ goes to $+\infty$, it is clear from the above conditions that the broken geodesic γ reduces to a single point, let's say $\gamma(t) = p$, $\forall t \in [0, 1]$. Moreover, for each $i = 0, \dots, N-1$, the parallel transport P_i coincides with the identity map and the last equation in (3.30) becomes

$$\exp_p^{-1}(p_0) + \exp_p^{-1}(p_1) + \cdots + \exp_p^{-1}(p_N) = 0,$$

which proves that p is in fact the Riemannian mean of the points p_i. $\square$

In figures 4-7, we illustrate the previous result for the particular cases when M is the Euclidean space $\mathbb{R}^2$ and the three-dimensional unit sphere S^2.

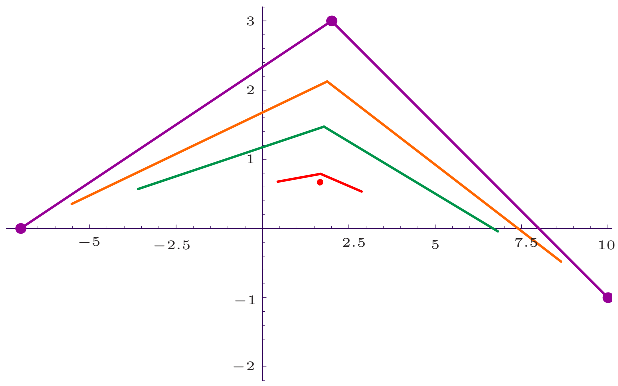

Figure 4: The data are the following: $q_0 = \big(-7, 0\big)$, $q_1 = (2, 3)$, $q_2 = (10, -1)$, $t_0 = 0$, $t_1 = \frac{1}{2}$ and $t_2 = 1$. The smoothing cubic splines were obtained for the following values of λ: $\lambda_1 = 10^{-5}$, $\lambda_2 = 10^{-1}$, $\lambda_3 = 10^{-0.5}$, $\lambda_4 = 3$ and $\lambda_5 = 10^3$.

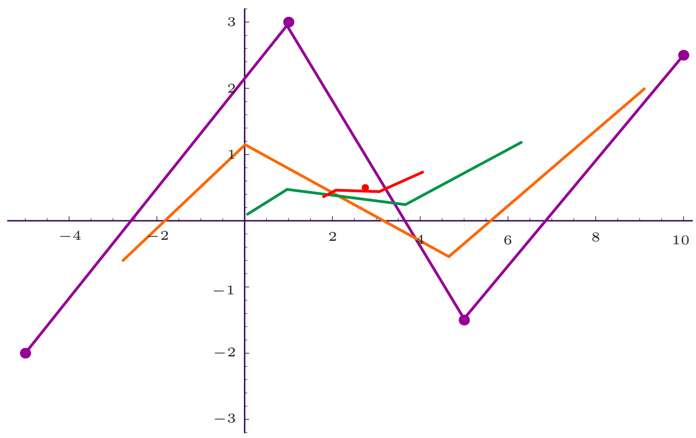

Figure 5: The data are the following: $q_0 = \big(-5, -2\big)$, $q_1 = (1, 3)$, $q_2 = \big(5, -\frac{3}{2}\big)$, $q_3 = \big(10, \frac{5}{2}\big)$, $t_0 = 0$, $t_1 = \frac{1}{8}$, $t_2 = \frac{1}{2}$ and $t_3 = 1$. The smoothing cubic splines were obtained for the following values of λ: $\lambda_1 = 10^{-3}$, $\lambda_2 = 10^{-1}$, $\lambda_3 = 0.7$, $\lambda_4 = 3$ and $\lambda_5 = 10^3$.

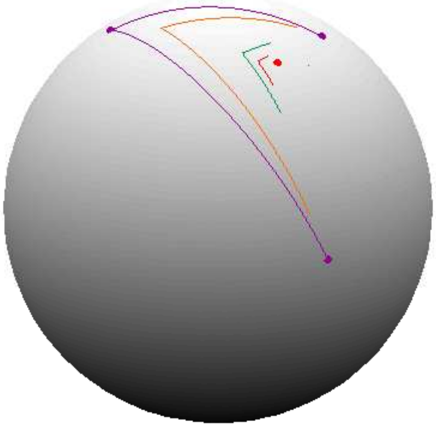

Figure 6: The data are the following: $q_0 = \left(\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}, \frac{\sqrt{3}}{2}\right)$, $q_1 = \left(-\frac{1}{4}, 0, \frac{\sqrt{3}}{2}\right)$, $q_2 = \left(1, 0, 0\right)$, $t_0 = 0$, $t_1 = \frac{1}{2}$ and $t_2 = 1$. The smoothing cubic splines were obtained for the following values of λ: $\lambda_1 = 10^{-3}$, $\lambda_2 = 10^{-1}$, $\lambda_3 = 1$, $\lambda_4 = 5$ and $\lambda_5 = 10^4$.

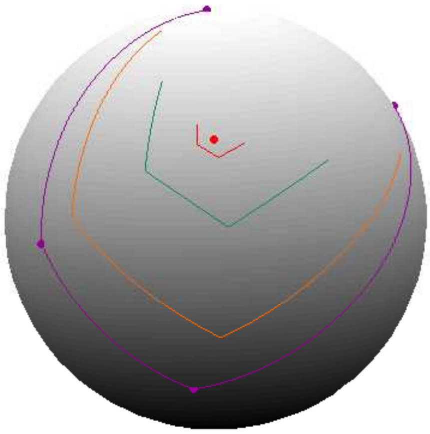

Figure 7: The data are the following: $q_0 = \left(0, 0, 1\right)$, $q_1 = (0, -1, 0)$, $q_2 = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{\sqrt{2}}{2}\right)$, $q_3 = \left(\frac{\sqrt{3}}{4}, \frac{3}{4}, \frac{1}{2}\right)$, $t_0 = 0$, $t_1 = \frac{1}{3}$, $t_2 = \frac{2}{3}$ and $t_3 = 1$. The smoothing cubic splines were obtained for the following values of λ: $\lambda_1 = 10^{-4}$, $\lambda_2 = 10^{-1}$, $\lambda_3 = 0.5$, $\lambda_4 = 3$ and $\lambda_5 = 10^4$.

4 Smoothing Splines and Least Squares Problems

In this section, we will finally establish the relationship between the smoothing geometric splines defined in the previous section and the solution of the classical least squares problem in Euclidean spaces. The results that we develop throughout this section, which generalize the classical results appearing in [23] for the particular case when $m = 2$ and $M = \mathbb{R}^n$, strong support our belief that the variational problem formulated at the very beginning of section 3 is the most natural way of generalizing the classical least squares problem to Riemannian manifolds.

Before we establish our main result in this section, we will briefly recall the classical least squares problem in Euclidean spaces.

4.1 Recalling the Classical Least Squares Problem

In the classical least squares problem, we are given a finite set of points in $\mathbb{R}^n$, $p_0, \ldots, p_N$, and a monotone increasing sequence of instants of time $t_0 < \cdots < t_N$, and the objective is to find a polynomial curve $t \mapsto \gamma(t) = a_0 + a_1 t + \cdots + a_{m-1} t^{m-1}$, with $m - 1 \leq N$, that minimizes the sum of the squared Euclidean distances from p_i to $\gamma(t_i)$. That is, that yields the minimum value for the functional

$$E(\gamma) = \sum_{i=0}^{N} \big\| p_i - \gamma(t_i) \big\|^2. \tag{4.31}$$

Although the classical literature only treats this problem for data in $\mathbb{R}$, its generalization to more general Euclidean spaces is straightforward.

It is easy to prove that the above problem has a unique solution γ, that is obtained by solving the following system of equations:

$$\begin{array}{l} \displaystyle\sum_{i=0}^{N} \gamma(t_i) = \sum_{i=0}^{N} p_i \\ \displaystyle\sum_{i=0}^{N} t_i \gamma(t_i) = \sum_{i=0}^{N} t_i p_i \\ \qquad\vdots \\ \displaystyle\sum_{i=0}^{N} t_i^{m-1} \gamma(t_i) = \sum_{i=0}^{N} t_i^{m-1} p_i \end{array}, \tag{4.32}$$

known in the literature as the normal equations [15].

4.2 Main Results

In what follows, we will assume that the Riemannian manifold M is the Euclidean space $\mathbb{R}^n$, endowed with the Riemannian metric induced by the Euclidean inner product.

Theorem 4.1. *When $M = \mathbb{R}^n$ and λ goes to $+\infty$ in conditions* (3.23)-(3.24) *of theorem* 3.1*, the smoothing splines converge to the polynomial of degree $m-1$ that is the solution of the classical least squares problem.*

Proof. For the case when M is the Euclidean space $\mathbb{R}^n$, the curvature tensor vanishes, the covariant derivative reduces to the usual derivative and therefore the differential equation (3.23) becomes simply

$$\frac{d^{2m}\gamma}{dt^{2m}} = 0. \tag{4.33}$$

The regularity conditions (3.24) take also the form

$$\begin{array}{l} \displaystyle\frac{d^k\gamma}{dt^k}(t_i^+) = \frac{d^k\gamma}{dt^k}(t_i^-) = 0, \qquad k = m, \ldots, 2m-2 \\ \displaystyle\frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_i^+) - \frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_i^-) = \frac{(-1)^m}{\lambda}\big(p_i - \gamma(t_i)\big) \end{array}, \tag{4.34}$$

for $i = 0, \ldots, N$.

Equation (4.33) can be integrate explicitly is each subinterval $\big[t_i, t_{i+1}\big]$. Let us consider

$$\gamma(t) = a_0^i + a_1^i t + \ldots + a_{2m-1}^i t^{2m-1}, \tag{4.35}$$

where $a_k^i \in \mathbb{R}^n$, for each $k = 0, \ldots, 2m-1$ and $i = 0, \ldots, N-1$.

Computing successively the derivatives of γ with respect to t, we get, for $t \in \big[t_i, t_{i+1}\big]$,

$$\begin{aligned}
&\frac{d\gamma}{dt}(t) = a_1^i + 2a_2^i t + \cdots + (2m-1)a_{2m-1}^i t^{2m-2}\\
&\frac{d^2\gamma}{dt^2}(t) = 2a_2^i + 3!a_3^i t + \cdots + (2m-1)(2m-2)a_{2m-1}^i t^{2m-3}\\
&\quad\vdots\\
&\frac{d^k\gamma}{dt^k}(t) = k!a_k^i + (k+1)!a_{k+1}^i t + \cdots + \frac{(2m-1)!}{(2m-k-1)!}a_{2m-1}^i t^{2m-k-1}\\
&\quad\vdots\\
&\frac{d^{2m-1}\gamma}{dt^{2m-1}}(t) = (2m-1)!a_{2m-1}^i
\end{aligned}.$$

Now, attending to the expression for the derivative of γ of order $2m-1$, it follows immediately that

$$\frac{d^{2m-1}\gamma}{dt^{2m-1}}\big(t_i^+\big) = \frac{d^{2m-1}\gamma}{dt^{2m-1}}\big(t_{i+1}^-\big), \tag{4.36}$$

where $i = 0, \ldots, N-1$.

Equality (4.36) can now be used to rewrite the last set of equations appearing in (4.34) as

$$\begin{aligned}
&\frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_0^+) = \frac{(-1)^m}{\lambda}\big(p_0 - \gamma(t_0)\big)\\
&\frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_1^+) - \frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_0^+) = \frac{(-1)^m}{\lambda}\big(p_1 - \gamma(t_1)\big)\\
&\qquad\qquad\vdots\\
&\frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_{N-1}^+) - \frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_{N-2}^+) = \frac{(-1)^m}{\lambda}\big(p_{N-1} - \gamma(t_{N-1})\big)\\
&\qquad\qquad\qquad - \frac{d^{2m-1}\gamma}{dt^{2m-1}}(t_{N-1}^+) = \frac{(-1)^m}{\lambda}\big(p_N - \gamma(t_N)\big)
\end{aligned}. \tag{4.37}$$

Adding up both terms of the above system of equations, we obtain

$$\frac{(-1)^m}{\lambda}\sum_{i=0}^{N}\Big(p_i - \gamma(t_i)\Big) = 0, \tag{4.38}$$

which is equivalent to the first equation of the normal equations (4.32).

According to the explicit form (4.35) of the curve γ in each subinterval $\big[t_i, t_{i+1}\big]$, we can also write the last equation of (4.34) as

$$(2m-1)!\Big(a_{2m-1}^i - a_{2m-1}^{i-1}\Big) = \frac{(-1)^m}{\lambda}\big(p_i - \gamma(t_i)\big). \tag{4.39}$$

The corresponding condition for the derivative of γ of order $2m-2$ can also be written as

$$(2m-2)!\Big(a^i_{2m-2} - a^{i-1}_{2m-2}\Big) = -(2m-1)!t_i\Big(a^i_{2m-1} - a^{i-1}_{2m-1}\Big). \tag{4.40}$$

Plugging equation (4.39) into (4.40) and then summing up both terms of the previous $N+1$ equations, we conclude that

$$\frac{(-1)^m}{\lambda}\sum_{i=0}^{N} t_i\Big(p_i - \gamma(t_i)\Big) = 0,$$

which is equivalent to the second equation of (4.32).

To complete the proof, we claim that for $l \in \{2, \dots, m\}$, the condition fulfilled by the derivative of γ of order $2m-l$ is equivalent to

$$(l-1)!(2m-l)!\Big(a^i_{2m-l} - a^{i-1}_{2m-l}\Big) = (-1)^{l-1}(2m-1)!t_i^{l-1}\Big(a^i_{2m-1} - a^{i-1}_{2m-1}\Big). \tag{4.41}$$

If the above condition holds, then plugging (4.36) into (4.41) and then summing up those $N+1$ equations, we get

$$\frac{(-1)^m}{\lambda}\sum_{i=0}^{N} t_i^{l-1}\Big(p_i - \gamma(t_i)\Big) = 0,$$

for $l = 2, \dots, m$.

When λ goes to $+\infty$, we have already proved in proposition 3.2 (b), that the smoothing spline γ approaches an Euclidean polynomial of degree $m-1$. On the other hand, since the above $m-1$ equations together with equation (4.38) are equivalent to the normal equations (4.32), that Euclidean polynomial is therefore the solution of the classical least squares problem.

Let us assume that condition (4.41) holds for $l \in \{2, \dots, m-1\}$ and let us prove that it still holds for $l+1$.

The condition appearing in (4.34) for the derivative of γ of order $2m-l-1$ can be written as

$$(2m-l-1)!\Big(a^i_{2m-l-1} - a^{i-1}_{2m-l-1}\Big) + (2m-l)!\, t_i \left(a^i_{2m-l} - a^{i-1}_{2m-l}\right) +$$
$$+ \cdots + \frac{(2m-2)!}{(l-1)!}\, t_i^{l-1}\left(a^i_{2m-2} - a^{i-1}_{2m-2}\right) + \frac{(2m-1)!}{l!}\, t_i^l \left(a^i_{2m-1} - a^{i-1}_{2m-1}\right) = 0.$$

Now, if we use the induction step (4.41), we obtain after some manipulations

$$\begin{aligned} l!(2m-l-1)!\Big(a^i_{2m-l-1} - a^{i-1}_{2m-l-1}\Big) &= \\ = -\sum_{j=1}^{l}(-1)^{l-j}\frac{l!}{(l-j)!j!}(2m-1)!\, t_i^l \left(a^i_{2m-1} - a^{i-1}_{2m-1}\right) \\ = (-1)^{l+1}(2m-1)!\, t_i^l \left(a^i_{2m-1} - a^{i-1}_{2m-1}\right)^1, \end{aligned}$$

which finishes the proof. ∎

For the particular case when $m = 2$, we conclude from the previous theorem that the straight line obtained by the described limiting process is indeed the solution of the corresponding classical least squares problem, thus also generalizing the results appearing in [23] and in [18].

We finish with some illustrations in the plane $\mathbb{R}^2$ of the results presented here where we can see that polynomials that are the solution of the classical least squares problems are obtained by this limiting process.

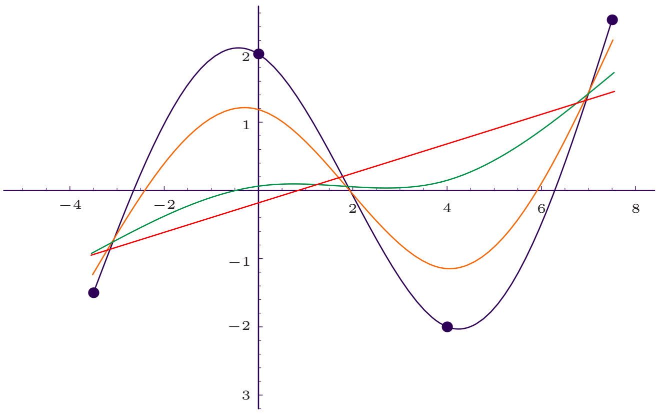

Figure 8: The data are the following: $q_0 = \left(-\frac{7}{2}, -\frac{3}{2}\right)$, $q_1 = (0, 2)$, $q_2 = (4, -2)$, $q_3 = \left(\frac{15}{2}, \frac{5}{2}\right)$, $t_0 = 0$, $t_1 = \frac{1}{3}$, $t_2 = \frac{2}{3}$ and $t_3 = 1$. The smoothing cubic splines were obtained for the following values of λ: $\lambda_1 = 10^{-5}$, $\lambda_2 = 10^{-3}$, $\lambda_3 = 10^{-2}$ and $\lambda_4 = 10$.

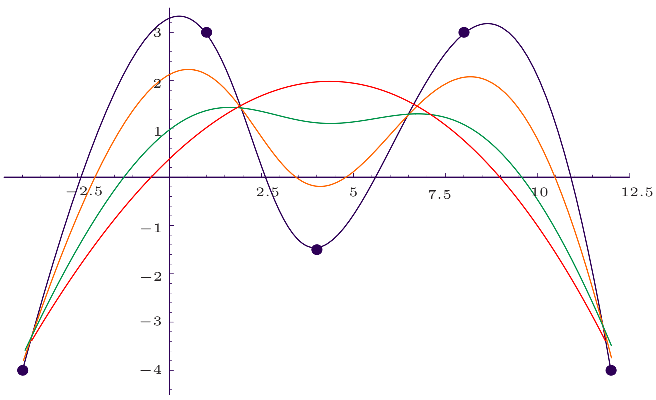

Figure 9: The data are the following: $q_0 = (-4, -4)$, $q_1 = (1, 3)$, $q_2 = \left(4, -\frac{3}{2}\right)$, $q_3 = (8, 3)$, $q_4 = (12, -4)$, $t_0 = 0$, $t_1 = \frac{1}{4}$, $t_2 = \frac{1}{2}$, $t_3 = \frac{3}{4}$ and $t_4 = 1$. The smoothing splines of degree 5 were obtained for the following values of λ: $\lambda_1 = 10^{-7}$, $\lambda_2 = 10^{-5}$, $\lambda_3 = 10^{-4.6}$ and $\lambda_4 = 10^3$.

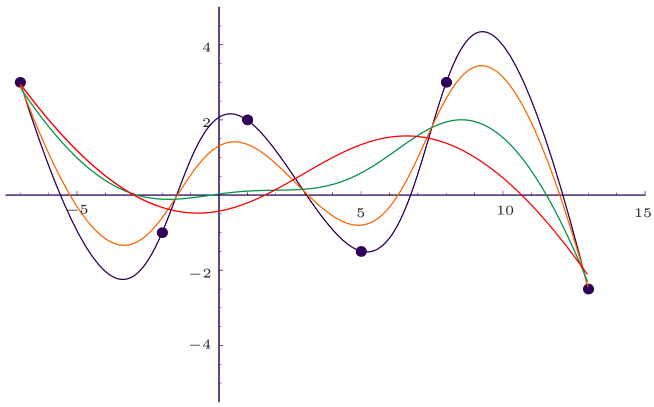

Figure 10: The data are the following: $q_0 = (-7, 3)$, $q_1 = (-2, -1)$, $q_2 = (1, 2)$, $q_3 = \left(5, -\frac{3}{2}\right)$, $q_4 = (8, 3)$, $q_5 = \left(13, -\frac{5}{2}\right)$, $t_0 = 0$, $t_1 = \frac{1}{5}$, $t_2 = \frac{2}{5}$, $t_3 = \frac{3}{5}$, $t_4 = \frac{4}{5}$ and $t_5 = 1$. The smoothing splines of degree 7 were obtained for the following values of λ: $\lambda_1 = 10^{-10}$, $\lambda_2 = 10^{-8}$, $\lambda_3 = 10^{-7}$ and $\lambda_4 = 10^{-4}$.

5 Conclusion

In this paper, we presented a generalization of high order classical least squares problems to more general Riemannian manifolds.

The formulation of the classical least squares problem given at the very beginning of section 4 has not a straightforward generalization to more general Riemannian manifolds. In fact, the non availability of explicit forms for the analogous to polynomial curves on manifolds was the main drawback to establish this generalization.

Nevertheless, the variational approach used to define such polynomial curves referred in subsection 2.2 enabled us to formulate in section 3 the variational problem $(\mathcal{P})$, depending on a smoothing parameter, and giving rise to what we call smoothing geometric splines.

These curves fit the given data and are obtained by piecing smoothly together segments of geometric polynomials. The Riemannian mean of the given points could also be obtained as a limiting process of those smoothing geometric splines, as it was proved in Theorem 3.3.

It was also possible to prove in Theorem 4.1, that when the smoothing parameter goes to infinity, the smoothing geometric curves approach a smooth curve that turns out to be the solution of the classical least squares problem for the particular case when the manifold reduces to an Euclidean space.

These facts were illustrated in the plane $\mathbb{R}^2$ and in the sphere S^2, in figures , using the software Matlab 7.1 and Mathematica 5.1.

References

[1] S. R. Buss and J. P. Fillmore, *Spherical Averages and Applications to Spherical Splines and Interpolation*, ACM Transactions on Graphics **2** (2001), no. 20, 95–126.

[2] M. Camarinha, *The Geometry of Cubic Polynomials on Riemannian Manifolds*, PhD Thesis, Departamento de Matemática, Universidade de Coimbra, Portugal, 1996.

[3] M. Camarinha, F. Silva Leite, and P. Crouch, *Splines of class C^k on Non-Euclidean Spaces*, IMA Journal of Mathematical Control and Information **12** (1995), 399–410.

[4] ______, *On the Geometry of Riemannian Cubic Polynomials*, Differential Geometry and its Applications (2001), no. 15, 107–135.

[5] J. M. Corcuera and W. Kendall, *Riemannian Barycentres and Geodesic Convexity*, Math. Proc. Camb. Phil. Soc. (1999), no. 127, 253–269.

[6] P. Crouch and F. Silva Leite, *Geometry and the Dynamic Interpolation Problem*, Proc. American Control Conference, Boston (1991), 1131–1137.

[7] ______, *The Dynamic Interpolation Problem: on Riemannian Manifolds, Lie Groups and Symmetric Spaces*, Journal of Dynamical and Control Systems **1** (1995), no. 2, 177–202.

[8] M. P. do Carmo, *Riemannian Geometry*, Mathematics: Theory and Applications, Birkäuser, 1992.

[9] G. Farin, *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, 1990.

[10] R. Giamgò, F. Giannoni, and P. Piccione, *An Analytical Theory for Riemannian Cubic Polynomials*, IMA J. Math. Control and Information **19** (2002), 445–460.

[11] K. Hüper and F. Silva Leite, *On the Geometry of Rolling and Interpolation Curves on S^n, SO_n and Grassmann Manifolds*, Journal of Dynamical and Control Systems **13** (2007), no. 4, 467–502.

[12] K. Hüper and J. H. Manton, *Numerical Methods to Compute the Karcher Mean of Points on the Special Orthogonal Group*, To appear.

[13] H. Karcher, *Riemannian Center of Mass and Mollifier Smoothing*, Communications on Pure and Applied Mathematics **30** (1977), 509–541.

[14] A. K. Krakowski, *Geometrical Methods of Inference*, PhD Thesis, Department of Mathematics and Statistics, The University of Western Australia, Australia, 2002.

[15] P. Lancaster and K. Salkauskas, *Curve and Surface Fitting*, Academic Press, 1990.

[16] J. M. Lee, *Riemannian Geometry: An Introduction to Curvature*, no. 176 in Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[17] L. Machado, *Least Squares Problems on Riemannian Manifolds*, PhD Thesis, Department of Mathematics, University of Coimbra, Portugal, 2006.

[18] L. Machado and F. Silva Leite, *Fitting Smooth Paths on Riemannian Manifolds*, International Journal of Applied Mathematics & Statistics **4** (2006), no. J06, 25–53.

[19] J. W. Milnor, *Morse Theory*, Princeton University Press, Princeton, New Jersey, 1963.

[20] M. Moakher, *A Differential Geometric Approach to the Arithmetic and Geometric Means of Operators in some Symmetric Spaces*, SIAM. J. Matrix Anal. Appl., **26** (2005), no. 3, 735–747.

[21] L. Noakes, G. Heinzinger, and B. Paden, *Cubic Splines on Curved Spaces*, IMA Journal of Mathematics Control and Information **6** (1989), 465–473.

[22] Tomasz Popiel and L. Noakes, *Higher Order Geodesics in Lie Groups*, Math. Control Signals Systems (2007), no. 19, 235–253.

[23] C. H. Reinsch, *Smoothing by Spline Functions*, Numerische Mathematik **10** (1967), 177–183.

[24] F. Silva Leite and P. Crouch, *Closed Forms for the Exponential Mapping on Matrix Lie Groups Based on Putzer's Method*, Journal of Mathematical Physics **40** (1999), 3561–3568.