Spherical Triangular Configurations with Invariant Geometric Mean

Luís Machado[a,b], Knut Hüper[c], Krzysztof Krakowski[d], Fátima Silva Leite[b,e]

[a]*Department of Mathematics, University of Trás-os-Montes and Alto Douro (UTAD), Quinta de Prados, Vila Real, 5001-801, Portugal*
[b]*Institute of Systems and Robotics, University of Coimbra, DEEC - Polo II, 3030-290, Coimbra, Portugal*
[c]*Institute of Mathematics, Julius-Maximilians-Universität Würzburg, Campus Hubland Nord Emil-Fischer-Straße 31, 97074, Würzburg, Germany*
[d]*Matematyczno-Przyrodniczy, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie (UKSW), Kazimierza Wýcickiego 1/3, 01-938, Warsaw, Poland*
[e]*Department of Mathematics, University of Coimbra, Largo D. Dinis, 3000-143, Coimbra, Portugal*

Abstract

The main objective is to characterize all configurations of three distinct points on the n-dimensional sphere that have the same Riemannian geometric mean and find efficient ways to compute such invariant. The regular case, when the points form the vertices of an equilateral spherical triangle, appears as the global minimum of an appropriate cost function. As a warm-up, and also to get more insight for the spherical case, we first develop our ideas for configurations in the Euclidean space $\mathbb{R}^n$. In both cases, the theoretical results are supported by numerical experiments and illustrated by meaningful plots.

Keywords: Geometric mean, Riemannian manifold, Riemannian gradient and Hessian, n-dimensional sphere, steepest descent, Newton's method.
2020 MSC: 15A10, 15A24, 49M05, 53C35

1. Introduction

In recent years there has been increasing interest in studying certain k-point configurations or arrangements on specific finite dimensional Riemannian manifolds, in particular configurations that fulfil certain geometric constraints. Among them are, for instance, those having their geometric mean in common or those maximizing the content of their convex hull. The first of these problems appears prominently in statistics on manifolds, where usually k points are given and one aims to find the geometric (Riemannian) mean or a closely related different type of weighted mean, cf. [3] and further references cited therein. The second is related to packing problems and as a consequence also to the task of designing codes with additional properties, say self-duality or fulfilling an additional optimality criterion. In this case, one looks for k points which satisfy certain geometric properties, cf. [8]. However, in both cases usually closed form solutions are extremely rare to find. So, a maximal amount of a priori (differential) geometric insight might be helpful for designing an efficient numerical procedure. Clearly, ordinary Euclidean geometry, along with the entire arsenal of linear algebra, is helpful simply because Euclidean space serves as a joyful playground for most of these problems. In such cases, closed-form solutions are indeed well-known, sometimes even for centuries.

Although, at first glance, the mathematics behind such goals is mainly based on differential geometric methodologies or insights, as well on purely algebraic grounds (other than the real numbers field), oftentimes one has to apply sophisticated numerical techniques, such as geometric integration, cf. [10], or geometric optimization, cf. [1].

We are generally interested in characterizing configurations of k distinct points in an n-dimensional Riemannian manifold that share the same geometric mean. Additionally, we aim to explore efficient methods for computing these configurations and their associated invariant.

Finding the geometric mean of data points on a Riemannian manifold has been extensively studied for quite some time. Typically, the geometric mean is the solution of an optimization problem, where the sum of the square geodesic distances to the data points is minimized. This approach has been used in specific manifolds, such as, S^n, the orthogonal group, the hyperbolic space, and the cone of positive symmetric matrices, cf., [4], [23], [6], [3], [18], [6], [5], [24], to name a few. For more differential geometric background with respect to existence and uniqueness of geometric means see [13] or [15]. In the literature, the geometric or Fréchet mean has, in some instances, been attributed to H. Karcher see, however, [14] (https://arxiv.org/abs/1407.2087).

Our general interest and objective is rather ambitious, since it requires more advanced backgrounds and time to mature ideas and achieve solid developments. To keep this paper in a manageable frame, for the moment we only consider the example of the standard sphere S^n embedded in the

Euclidean space $\mathbb{R}^{n+1}$, with $k = 3$, but using new ideas rather than just minimizing the sum of the squared geodesic distances. Some of these new ideas, also emerged from the fact that one often knows a simple formula for the mean of k points that form a regular Riemannian geodesic k-gon, e.g. for $k = 3$ an equilateral geodesic triangle. In order to get some insight, we nevertheless, even start by applying this new approach to Euclidean n-space. Generalizations to other symmetric spaces are already in preparation and will appear in forthcoming publications.

The paper is organized as follows. After introducing the necessary notations, our problem statement is presented in Section 3, where we consider, as warming up, points in the Euclidean space $\mathbb{R}^n$. Then, in Section 4, we transfer all the ideas and procedures to the n-dimensional spherical case. The geometry of the manifold consisting of all configurations of points that have the same Riemannian geometric mean is studied in detail. Both sections, 3 and 4, also contain explicit calculations, a detailed characterization of the critical point sets of appropriate cost functions, their Riemannian gradients and Hessians, followed by the classification of the critical points. In particular, the equilateral triangle configurations arise as global minimum of those cost functions.

Using several routines from MATLAB toolboxes, the steepest descent and quasi-Newton algorithms on manifolds have been implemented to corroborate the theoretical outcomes. These algorithms turned out to be easy to implement, offering high accuracy and precision even when handling spherical data. To enrich the paper, meaningful plots illustrating our results are also included.

2. Notations

These are some of the notations used throughout the paper.

M, N	smooth manifolds
T_xM	tangent space of M at $x \in M$
$DF(x)\colon T_xM \to T_{F(x)}N$	tangent map (or differential) of $F\colon M \to N$ at x
$T_x^{\perp}M$	normal space of M at $x \in M$
∇F	gradient of the function F
H_F	Hessian matrix of the function F
S^n	n-dimensional unit sphere
$\mathcal{H}^n$	open hemisphere of S^n
$\lVert . \rVert$	Euclidean norm
$\cos^{-1}(x^\top y)$	$\arccos(x^\top y)$
$\cos^{-2}(x^\top y)$	$\arccos^2(x^\top y)$
$\ker(X)$	kernel of a matrix X
$RS(X)$	row space of a matrix X
$CS(X)$	column space of a matrix X
$rref(X)$	reduced row echelon form of a matrix X
X^+	Moore-Penrose inverse of a matrix X

3. Problem Statement, Warming up in $\mathbb{R}^n$

Given three distinct points x_0, x_1, x_2 in $\mathbb{R}^n$, find all configurations of three points $\{p_0, p_1, p_2\} \subset \mathbb{R}^n$ having the same geometric mean q as the given ones, i.e.,

$$q = \tfrac{1}{3}(p_0 + p_1 + p_2) = \tfrac{1}{3}(x_0 + x_1 + x_2). \tag{1}$$

In particular, we are also looking for three points that form the vertices of a regular 3-gon, i.e., an equilateral triangular having q as the center of its circumscribed circle.

Recall that q is the unique solution of the minimization problem

$$\min_{x \in \mathbb{R}^n} \left(\|p_0 - x\|^2 + \|p_1 - x\|^2 + \|p_2 - x\|^2\right). \tag{2}$$

Here it seems we are dealing with a catch-22 as we do not need an equilateral triangle to verify formula (1), but we are currently only warming up for the much more complicated spherical case, where q is in general given only implicitly as the unique global minimum of a smooth cost.

Without loss of generality we fix one of the three points, say $x_0 =: p_0$, and consider the smooth manifold

$$M = \big\{(p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^n \;:\; p_1 + p_2 = 3q - p_0\big\}, \tag{3}$$

which is clearly an n-dimensional affine subspace. The tangent and the normal spaces of M at $(p_1, p_2) \in M$ are given, respectively, by

$$T_{(p_1,p_2)}M := \big\{(v, -v) \mid v \in \mathbb{R}^n\big\}, \quad T^{\perp}_{(p_1,p_2)}M := \big\{(v, v) \mid v \in \mathbb{R}^n\big\}. \tag{4}$$

Any vector $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n$ can be decomposed in a unique way as

$$(u, v) = \tfrac{1}{2}\big(u - v, v - u\big) + \tfrac{1}{2}\big(u + v, u + v\big), \tag{5}$$

where $\frac{1}{2}\big(u - v, v - u\big) \in T_{(p_1,p_2)}M$ and $\frac{1}{2}\big(u + v, u + v\big) \in T^{\perp}_{(p_1,p_2)}M$.

In the sequel we analyze the smooth cost function

$$\begin{aligned} F\colon M &\longrightarrow \mathbb{R}, \\ (p_1, p_2) &\longmapsto \tfrac{1}{4}\big(\|p_0 - p_1\|^2 - \|p_0 - p_2\|^2\big)^2 + \tfrac{1}{4}\big(\|p_1 - p_0\|^2 - \|p_1 - p_2\|^2\big)^2 \\ &\qquad + \tfrac{1}{4}\big(\|p_2 - p_0\|^2 - \|p_2 - p_1\|^2\big)^2. \end{aligned} \tag{6}$$

Clearly, the global minimum value of F equals 0 and it is attained exactly if the triple (p_0, p_1, p_2) describes an equilateral triangle in $\mathbb{R}^n$ or is the degenerated case when all the points coincide. One of our objectives is to minimize F to end up with one of these equilateral triangles.

To simplify notations, define

$$\begin{aligned} A &:= \|p_0 - p_1\|^2 - \|p_0 - p_2\|^2; \\ B &:= \|p_1 - p_0\|^2 - \|p_1 - p_2\|^2; \\ C &:= \|p_2 - p_0\|^2 - \|p_2 - p_1\|^2 = B - A. \end{aligned} \tag{7}$$

Consider the following smooth cost function

$$\begin{aligned} F\colon M &\longrightarrow \mathbb{R}, \\ (p_1, p_2) &\longmapsto \tfrac{1}{4}\big(A^2 + B^2 + C^2\big) = \tfrac{1}{2}(A^2 + B^2 - AB). \end{aligned} \tag{8}$$

Theorem 1. *Every critical point $(p_1, p_2) \in M$ of the function F defined by (8) fulfills one of the following conditions.*

1. $p_0 = p_1 = p_2 = q$;

2. $p_1 = p_2 = \frac{3q-p_0}{2}$;

3. p_0, p_1, p_2 *form an equilateral triangle.*

Proof. The critical points are the points (p_1, p_2) such that $\mathrm{D}F(p_1, p_2)(v, -v) = 0$, for all $v \in \mathbb{R}^n$, where $\mathrm{D}F$ stands for the differential of F. Since

$$\begin{aligned} \mathrm{D}F(p_1, p_2)(v, -v) &= (2A - B)\big(\langle p_1 - p_0, v\rangle + \langle p_2 - p_0, v\rangle\big) \\ &\quad + (2B - A)\big(\langle p_1 - p_0, v\rangle + \langle p_2 - p_1, 2v\rangle\big) \\ &= 3\langle A(p_1 - p_0) + B(p_2 - p_1), v\rangle, \end{aligned} \tag{9}$$

the critical points (p_1, p_2) are solutions of

$$A(p_1 - p_0) + B(p_2 - p_1) = 0. \tag{10}$$

Let us consider the following cases.

Case 1. $p_1 - p_0$ and $p_2 - p_1$ are linearly dependent.

Case 1.1. $p_1 - p_0 = 0$ and $p_2 - p_1 = 0$.
We get the trivial case $p_0 = p_1 = p_2 = q$.

Case 1.2. $p_1 - p_0 \neq 0$ and $p_2 - p_1 = 0$.
From the constraint $p_0 + p_1 + p_2 = 3q$, we immediately get $p_1 = p_2 = \frac{3q-p_0}{2}$. Note that in this case $A = 0$, so equation (10) is satisfied.

Case 1.3. $p_1 - p_0 = 0$ and $p_2 - p_1 \neq 0$.
Using the definition of B, this implies that $B \neq 0$. But on the other hand these conditions, together with (10), imply that $B = 0$. So, this case gives no critical points.

Case 1.4. $p_1 - p_0 \neq 0$ and $p_2 - p_1 \neq 0$.
In this case, there exists $\lambda \in \mathbb{R}$ ($\lambda \neq 0$) such that $p_1 - p_0 = \lambda(p_1 - p_2)$ and so the equation (10) is satisfied only when $B = \lambda A$. Using the fact that $p_0 - p_2 = (p_0 - p_1) + (p_1 - p_2) = (1 - \lambda)(p_1 - p_2)$, and replacing in the expressions of A and B, we get

$$A = (2\lambda - 1)\|p_1 - p_2\|^2, \quad B = (\lambda^2 - 1)\|p_1 - p_2\|^2. \tag{11}$$

But the condition $B = \lambda A$ implies that $\lambda^2 - \lambda + 1 = 0$, which has no real solutions. So, this case gives no critical points.

Case 2. $p_1 - p_0$ and $p_2 - p_1$ are linearly independent.
In this case $A = B = 0$ or, equivalently,

$$\|p_0 - p_1\|^2 = \|p_0 - p_2\|^2 = \|p_1 - p_2\|^2. \tag{12}$$

So, p_0, p_1, p_2 form the vertices of an equilateral triangle.

This completes the proof.

□

Remark 1. *Notice that the cost function F vanishes at the critical points corresponding to 1. and 3. in the previous theorem, while for the other case we have $A = 0$, $B = \|p_1 - p_0\|^2 > 0$, and so, $F(p_1, p_2) = \frac{B^2}{2} > 0$.*

3.1. Gradients and Hessians

The simplest way to compute Riemannian gradients and Riemannian Hessians for a function F defined on a Riemannian manifold M is by exploiting well known formulas for the Levi-Civita connection ∇. In particular, for the latter

$$\operatorname{Hess} F(X, Y) = \nabla_X(\nabla_Y F) - \mathrm{D}F(\nabla_X Y), \tag{13}$$

where X and Y are vector fields in M (see, for instance, pages 343-344 in [17]). There are two situations when the second summand in (13) vanishes. Either the Hessian is evaluated at a critical point p and $\mathrm{D}F(p) = 0$, or one considers the representation of the Hessian along a geodesic γ, in which case, $X = Y = \dot\gamma$ and consequently $\nabla_{\dot\gamma}\dot\gamma = 0$. In this paper, we only need to evaluate Riemannian Hessians in these two situations, and since we only consider submanifolds embedded in Euclidean spaces, the Riemannian Hessian coincides with the tangent space projection of the Euclidean Hessian.

We now consider gradients and Hessians of the function F defined by (8). In a straightforward way we extend F uniquely from M to a smooth function $\widehat{F}$ on the embedding space $\mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^{2n}$, compute the Euclidean gradient of $\widehat{F}$ and project it back to TM orthogonally to end up with the Riemannian gradient ∇F of F on M. The symbol ∇ used in(13) for the Levi-Civita connection will no longer be used later. So, our notation for the Riemannian gradient of a function will not be a source of confusion.

For any $(p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^n$ and $(v_1, v_2) \in T_{(p_1,p_2)}(\mathbb{R}^n \times \mathbb{R}^n) \cong \mathbb{R}^n \times \mathbb{R}^n$ we have

$$\mathrm{D}\widehat{F}(p_1, p_2)(v_1, v_2) = \langle \nabla \widehat{F}(p_1, p_2), (v_1, v_2) \rangle. \tag{14}$$

Since

$$
\begin{aligned}
\mathrm{D}\widehat{F}(p_1,p_2)(v_1,v_2) &= (2A-B)\big(\langle p_1-p_0,v_1\rangle + \langle p_0-p_2,v_2\rangle\big) \\
&\quad + (2B-A)\big(\langle p_2-p_0,v_1\rangle + \langle p_1-p_2,v_2\rangle\big) \\
&= \langle (2A-B)(p_1-p_0) + (2B-A)(p_2-p_0), v_1\rangle \\
&\quad + \langle (2A-B)(p_0-p_2) + (2B-A)(p_1-p_2), v_2\rangle,
\end{aligned}
\tag{15}
$$

the Euclidean gradient is given by

$$
\nabla\widehat{F}(p_1,p_2) = \begin{bmatrix} (2A-B)(p_1-p_0) + (2B-A)(p_2-p_0) \\ (2A-B)(p_0-p_2) + (2B-A)(p_1-p_2) \end{bmatrix}. \tag{16}
$$

Consequently, from (5) we get for the Riemannian gradient

$$
\nabla F(p_1,p_2) = P^{\perp}_{(p_1,p_2)} \nabla\widehat{F}(p_1,p_2) \in T_{(p_1,p_2)}M, \tag{17}
$$

where

$$
P^{\perp}_{(p_1,p_2)} = \frac{1}{2}\begin{bmatrix} I_n & -I_n \\ -I_n & I_n \end{bmatrix}. \tag{18}
$$

Some straightforward computations show that

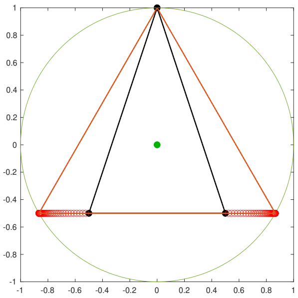
$$
\nabla F(p_1,p_2) = \frac{3}{2}\begin{bmatrix} A(p_1-p_0) + B(p_2-p_1) \\ -\big(A(p_1-p_0) + B(p_2-p_1)\big) \end{bmatrix}. \tag{19}
$$

The Riemannian gradient can now be used to implement the steepest descent method (Algorithm 1). Figure 1 illustrates this method for points in $\mathbb{R}^2$ with fixed $p_0 = (0,1)$. In the three situations shown on the left hand side, the points p_1 and p_2 (whose coordinates are given in each caption) converge to the minimum of the cost function, forming with p_0 the vertices of an equilateral triangle. The circumscribed circle observed in each picture is centered at the geometric mean, a property only shared by equilateral triangles. The graphs on the right hand side show how the distances between pairs of points (p_1,p_2) in successive iterations evolve. As expected, these distances converge linearly to zero.

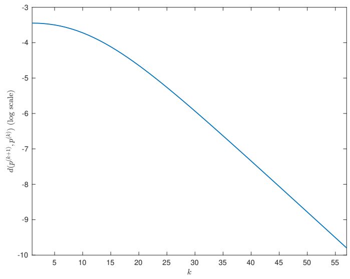
Algorithm 1: Steepest descent with Armijo line search

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)})$ and tolerance tol
Output: Stationary point $p^* = (p_1^*, p_2^*)$

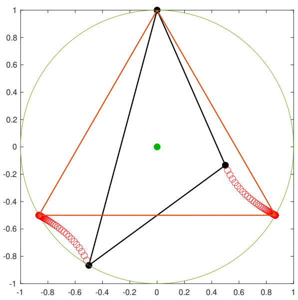
for $k = 0, 1, \ldots$ **do**
  Set $d^{(k)} = -\nabla F(p^{(k)})$;
  Determine the step length α_k according to Armijo rule;
  Set $p^{(k+1)} = p^{(k)} + \alpha_k d^{(k)}$;
  Stop if $F(p^{(k)}) < \mathsf{tol}$ or $\|\nabla F(p^{(k)})\| < \mathsf{tol}$
end



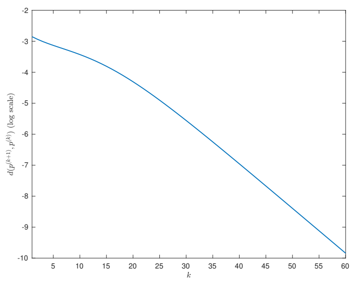
(a) $p_1 = (\frac{1}{2}, -\frac{1}{2})$, $p_2 = (-\frac{1}{2}, -\frac{1}{2})$



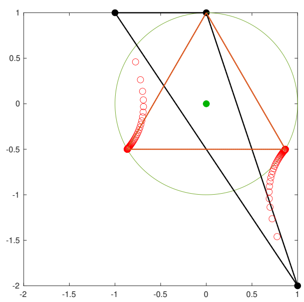
(b) Iteration k versus $d(p^{(k+1)}, p^{(k)})$



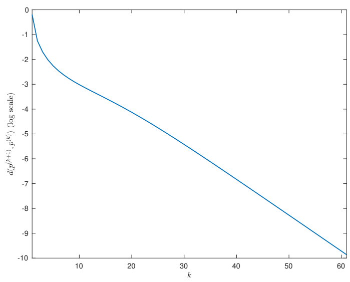
(c) $p_1 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$, $p_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2} - 1)$



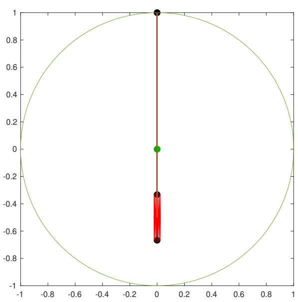
(d) Iteration k versus $d(p^{(k+1)}, p^{(k)})$



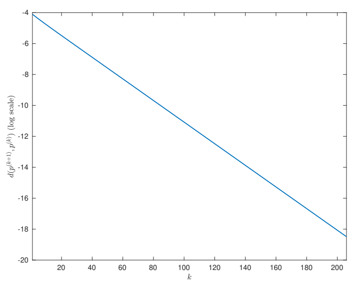
(e) $p_1 = (-1, 1)$, $p_2 = (1, -2)$



(f) Iteration k versus $d(p^{(k)}, p^{(k+1)})$



(g) $p_1 = (0, -\frac{1}{3})$, $p_2 = (0, -\frac{2}{3})$



(h) Iteration k versus $d(p^{(k)}, p^{(k+1)})$

Figure 1: Plots obtained using Algorithm 1.

Now, from the Euclidean gradient given in (16), we can proceed with the Hessian. The matrix representation of the Hessian of a function F will be denoted by H_F. In order to compute second derivatives, notice that

$$
\begin{aligned}
2A - B &= \|p_0 - p_1\|^2 - 2\|p_0 - p_2\|^2 + \|p_1 - p_2\|^2 \\
2B - A &= \|p_0 - p_1\|^2 - 2\|p_1 - p_2\|^2 + \|p_0 - p_2\|^2
\end{aligned}
$$

and we compute the matrix representation of the Hessian of $\widehat{F}$ as follows

$$
\begin{aligned}
\mathrm{D}_1^2\widehat{F}(p_1,p_2) =& 2\big((p_1 - p_0)(p_1 - p_0 + p_1 - p_2)^\top + (p_2 - p_0)(p_2 - p_0 + p_2 - p_1)^\top\big) \\
&+ (2A - B)I_n, \\
=& 2\big((p_1 - p_0)(p_1 - p_0)^\top + (p_2 - p_0)(p_2 - p_0)^\top + (p_2 - p_1)(p_2 - p_1)^\top\big) \\
&+ (2A - B)I_n, \\
\mathrm{D}_{12}\widehat{F}(p_1,p_2) =& 2\big((p_1 - p_0)(p_0 - p_1 + p_0 - p_2)^\top + (p_2 - p_0)(p_1 - p_2 + p_1 - p_0)^\top\big) \\
&+ (2B - A)I_n \\
=& 2\big((p_1 - p_0)(p_0 - p_2)^\top + (p_1 - p_2)(p_0 - p_1)^\top + (p_2 - p_0)(p_1 - p_2)^\top\big) \\
&+ (2B - A)I_n \\
=& \big(\mathrm{D}_{21}\widehat{F}(p_1,p_2)\big)^\top, \\
\mathrm{D}_2^2\widehat{F}(p_1,p_2) =& 2\big((p_0 - p_2)(p_0 - p_1 + p_0 - p_2)^\top + (p_1 - p_2)(p_1 - p_2 + p_1 - p_0)^\top \\
&- (A + B)I_n \\
=& 2\big((p_1 - p_0)(p_1 - p_0)^\top + (p_2 - p_0)(p_2 - p_0)^\top + (p_1 - p_2)(p_1 - p_2)^\top\big) \\
&- (A + B)I_n.
\end{aligned}
\tag{20}
$$

So,

$$
H_{\widehat{F}}(p_1,p_2) = \begin{bmatrix} \mathrm{D}_1^2\widehat{F}(p_1,p_2) & \mathrm{D}_{12}\widehat{F}(p_1,p_2) \\ \big(\mathrm{D}_{12}\widehat{F}(p_1,p_2)\big)^\top & \mathrm{D}_2^2\widehat{F}(p_1,p_2) \end{bmatrix}. \tag{21}
$$

Furthermore, being the Riemannian Hessian the restriction to the tangent space of the Euclidean Hessian, we can write

$$
H_F(p_1,p_2) = H_{\widehat{F}}(p_1,p_2)\big|_{T_{(p_1,p_2)}M} = P^\perp_{(p_1,p_2)} H_{\widehat{F}}(p_1,p_2) P^\perp_{(p_1,p_2)}. \tag{22}
$$

Here an important remark is in order. $H_F(p_1,p_2)$ is considered here in coordinates of the embedding space $\mathbb{R}^n \times \mathbb{R}^n$, the space of point pairs (p_1,p_2).

It, however, defines a symmetric quadratic form on the subspace $T_{(p_1,p_2)}M \subset \mathbb{R}^n \times \mathbb{R}^n$ which has codimension n.

The matrix representation of the Riemannian Hessian is a $2n \times 2n$ block matrix with the structure

$$H_F(p_1, p_2) = \begin{bmatrix} X & -X \\ -X & X \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \otimes X, \tag{23}$$

where X is written in terms of A and B as

$$\begin{aligned} X = \frac{3}{4}\big(&2(p_1 - p_0)(p_1 - p_0)^\top \\ &+ 2(p_2 - p_0)(p_2 - p_0)^\top + 2(p_1 - p_2)(p_1 - p_2)^\top + (A - 2B)I\big), \end{aligned} \tag{24}$$

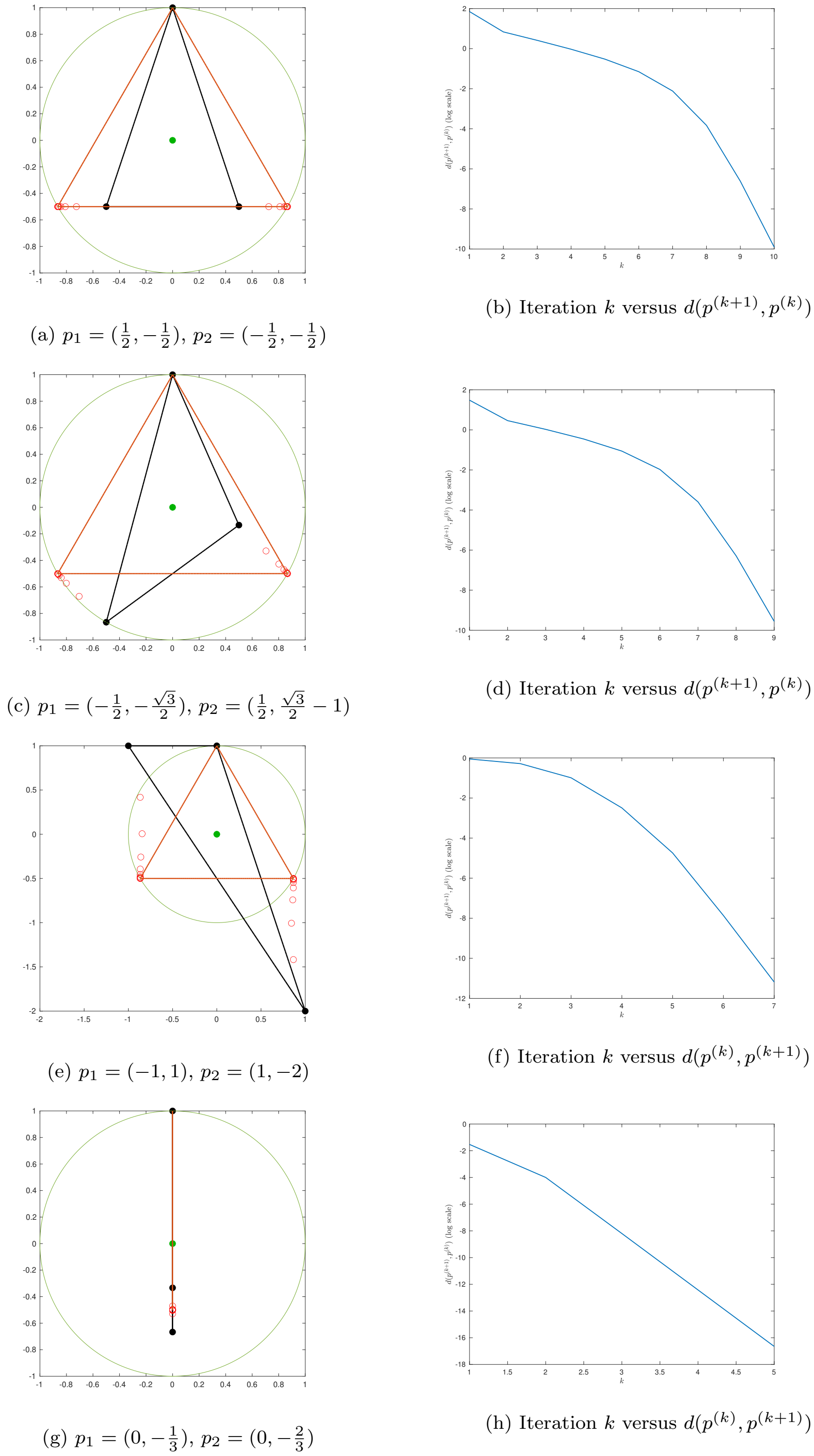
and is clearly symmetric and singular.

Making use of the Riemannian Hessian, we now apply a quasi-Newton method to improve the convergence speed (Algorithm 4). Figure 2 illustrates this method for the data already used to implement Algorithm 1. As expected, the distances between pairs of points (p_1, p_2) in successive iterations converge quadratically to zero, showing that the quasi-Newton method is faster than the steepest descent method.

Algorithm 2: Quasi-Newton method

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)})$, $\lambda > 0$ and tolerance tol
Output: Stationary point $p^* = (p_1^*, p_2^*)$

for $k = 0, 1, \dots$ **do**
  Set $B^{(k)} = H_F(p^{(k)}) + \lambda I_n$;
  Solve $d^{(k)}$ from $B^{(k)} d^{(k)} = -\nabla F(p^{(k)})$;
  Update $p^{(k+1)} = p^{(k)} + d^{(k)}$;
  Stop if $F(p^{(k)}) < \mathsf{tol}$ or $\|\nabla F(p^{(k)})\| < \mathsf{tol}$
end



(a) $p_1 = (\frac{1}{2}, -\frac{1}{2})$, $p_2 = (-\frac{1}{2}, -\frac{1}{2})$

(b) Iteration k versus $d(p^{(k+1)}, p^{(k)})$

(c) $p_1 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$, $p_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2} - 1)$

(d) Iteration k versus $d(p^{(k+1)}, p^{(k)})$

(e) $p_1 = (-1, 1)$, $p_2 = (1, -2)$

(f) Iteration k versus $d(p^{(k)}, p^{(k+1)})$

(g) $p_1 = (0, -\frac{1}{3})$, $p_2 = (0, -\frac{2}{3})$

(h) Iteration k versus $d(p^{(k)}, p^{(k+1)})$

Figure 2: Plots obtained using Algorithm 4.

3.2. Classification of the critical points

At the critical points given in Theorem 1, the formulas (21) and(22) simplify considerably, either by collinearity of the triple $\{p_0, p_1, p_2\}$ or by equilaterality. The latter suggests to deal with the triple in $\mathbb{R}^2$, more specifically in the plane spanned by this triple.

For convenience, we may represent pairs of points in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ by (v_1, v_2) or by a column matrix $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

Theorem 2. *The critical points of the function F on M are classified as:*

1. *when $p_1 = p_2 = \frac{3q-p_0}{2}$, the critical point (p_1, p_2) is a saddle point;*
2. *when p_0, p_1, p_2 form an equilateral triangle, the critical point (p_1, p_2) is a global minimum.*

Proof. In order to show that the critical point (p_1, p_1) is a saddle point, first note that in this case of collinearity $A = 0$, $B = \|p_1 - p_0\|^2$, and consequently

$$X = \frac{3}{2}\big(2(p_1 - p_0)(p_1 - p_0)^T - \|p_1 - p_0\|^2 I\big).$$

First consider the vector $v = (v_1, -v_1)$, where $v_1 = p_1 - p_0$. In this case,

$$\begin{aligned} \begin{bmatrix} v_1^\top & -v_1^\top \end{bmatrix} H_F(p_1, p_2) \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix} &= \begin{bmatrix} v_1^\top & -v_1^\top \end{bmatrix} \begin{bmatrix} X & -X \\ -X & X \end{bmatrix} \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix} \\ &= 4\, v_1^\top X v_1 = 6\, \|p_1 - p_0\|^4 > 0. \end{aligned} \tag{25}$$

Now, consider a vector $v = (v_1, v_2)$ such that $(v_1 - v_2)^\top (p_1 - p_0) = 0$. In this case, we get

$$\begin{aligned} \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix} H_F(p_1, p_2) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} &= \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix} \begin{bmatrix} X & -X \\ -X & X \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= (v_1 - v_2)^\top X (v_1 - v_2) \\ &= -\tfrac{3}{2}\|p_1 - p_0\|^2 \|v_1 - v_2\|^2 < 0. \end{aligned} \tag{26}$$

This proves that we are in the presence of a saddle point.

In case of equilaterality, $A = B = 0$ and X is the rank two matrix

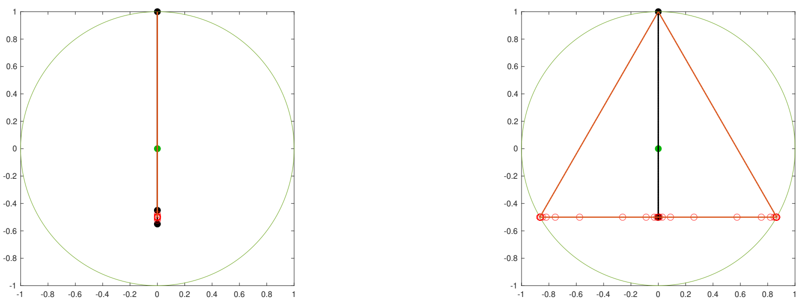
$$X = \frac{3}{2}\big((p_1 - p_0)(p_1 - p_0)^\top + (p_2 - p_0)(p_2 - p_0)^\top + (p_1 - p_2)(p_1 - p_2)^\top\big). \tag{27}$$

For convenience, denote the critical point by $(p_1^\star, p_2^\star)$. Then, for any vector $v = (v_1, v_2) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$,

$$\begin{aligned}\begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix} H_F(p_1^\star, p_2^\star) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} =& \tfrac{3}{2}\Big[\big((v_1 - v_2)^\top (p_1^\star - p_0)\big)^2 \\ &+ \big((v_1 - v_2)^\top (p_2^\star - p_0)\big)^2 \\ &+ \big((v_1 - v_2)^\top (p_1^\star - p_2^\star)\big)^2\Big] \geq 0,\end{aligned} \tag{28}$$

meaning that the Riemannian Hessian at $(p_1^\star, p_2^\star)$ is positive semidefinite. This, combined with the observation that $F(p_1, p_2) \geq F(p_1^\star, p_2^\star) = 0$, for all $(p_1, p_2) \in M$, leads to the conclusion that the cost function F attains its global minimum at $(p_1^\star, p_2^\star)$. □

Figure 3 shows the behavior of the points (p_1, p_2) in a neighborhood of a saddle point. If p_1 and p_2 are aligned with p_0, then they converge to the saddle point. Otherwise, they converge to the minimum of the cost function (the vertices of an equilateral triangle).



(a) $p_1 = (0, -0.55)$, $p_2 = (0, -0.45)$

(b) $p_1 = (-0.001, -0.499)$, $p_2 = (0.001, -0.501)$

Figure 3: Behavior in a neighborhood of a saddle point.

4. The Spherical Case

4.1. Some background

In preparation for the main results in this section, we first recall some important facts that will be used in this section, referring to [19] for basic concepts of Riemannian geometry.

Consider the unit sphere

$$S^n = \{p \in \mathbb{R}^{n+1} \mid p^\top p = 1\}, \tag{29}$$

equipped with the Riemannian metric induced by the Euclidean inner product in $\mathbb{R}^{n+1}$. Its tangent and normal space at $p \in S^n$ are, respectively,

$$T_pS^n = \{v \in \mathbb{R}^{n+1} \mid v^\top p = 0\}, \quad T_p^\perp S^n = \operatorname{span}(p). \tag{30}$$

If $p \in S^n$ and $v \in T_pS^n$, the unique minimal geodesic with $\gamma(0) = p$, $\dot\gamma(0) = v$ is given by

$$\gamma(t) = \cos(t\|v\|)p + \tfrac{\sin(t\|v\|)}{\|v\|}v. \tag{31}$$

If $p, q \in S^n$ with $p \neq \pm q$, the unique minimal geodesic satisfying $\gamma(0) = p$, $\gamma(1) = q$, purely expressed by p and q only, is given by (see, for instance [12]),

$$\gamma(t) = \tfrac{\sin((1-t)\|v\|)}{\sin\|v\|}p + \tfrac{\sin(t\|v\|)}{\sin\|v\|}q, \quad \text{with } \|v\| = \arccos(q^\top p). \tag{32}$$

The geodesic distance between two points $p, q \in S^n$ is exactly equal to the norm of the velocity vector that takes p to q, i.e.,

$$d(p, q) = \arccos(q^\top p). \tag{33}$$

For $x \in \mathbb{R}^{n+1}\backslash\{0\}$, the orthogonal projection operator is defined by

$$P_x^\perp \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \quad y \mapsto \big(I - \tfrac{xx^\top}{x^\top x}\big)y, \tag{34}$$

and the associated reflection operator by

$$R_x \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \quad R_x := \operatorname{id} - 2P_x^\perp. \tag{35}$$

The latter is an orthogonal linear transformation, thus preserving the Euclidean metric.

When $x = p \in S^n$,

$$P_p^\perp \colon \mathbb{R}^{n+1} \to T_pS^n, \quad y \mapsto \big(I - pp^\top\big)y, \tag{36}$$

and

$$R_p \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \quad y \mapsto \big(-I + 2pp^\top\big)y. \tag{37}$$

Notice that

$$\ker(P_p^\perp) = T_p^\perp S^n, \quad \operatorname{rank}(P_p^\perp) = n,$$

and

$$R_p|_{T_pS^n} = -\operatorname{id}, \quad R_p|_{T_p^\perp S^n} = \operatorname{id}.$$

4.2. Problem Statement on the sphere S^n

Given three distinct points x_0, x_1, x_2 contained in an open hemisphere $\mathcal{H}^n$ of S^n, find all configurations of points $\{p_0, p_1, p_2\}$ in that hemisphere that have the same Riemannian geometric mean $q \in \mathcal{H}^n$, in particular those that form the vertices of an equilateral spherical triangle centered at q. Without loss of generality we keep the assumption $p_0 := x_0$.

It is well known that the Riemannian mean minimizes the sum of squared geodesic distances to a given set of points. In our case, consider the smooth function

$$\Phi\colon \mathcal{H}^n \subset S^n \longrightarrow \mathbb{R}, \quad x \longmapsto \Phi(x) = \sum_{i=0}^{2} d^2(x_i, x) = \sum_{i=0}^{2} \arccos^2(x^\top x_i), \tag{38}$$

with tangent map

$$\begin{aligned} &\mathrm{D}\,\Phi(x)\colon T_x\mathcal{H}^n \to T_{\Phi(x)}\mathbb{R} \cong \mathbb{R}, \\ &\qquad h \mapsto 2\sum_{i=0}^{2} \underbrace{\arccos(x^\top x_i)}_{=:\xi_i} \mathrm{D}\arccos(x^\top x_i)h = -2\sum_{i=0}^{2} \tfrac{\xi_i}{\sin\xi_i}(x_i^\top h). \end{aligned} \tag{39}$$

[7], showed that Φ is strictly convex and therefore has exactly one global minimum, the geometric Riemannian mean, [13], [15], [16].

It is well known that the Riemannian mean is the solution of $\mathrm{D}\,\Phi(x)(h) = 0$, $\forall h \in T_x\mathcal{H}^n$. From now on, q will denote the Riemannian mean of the points x_i, which is defined implicitly by

$$\sum_{i=0}^{2} \tfrac{\xi_i}{\sin\xi_i}(I_{n+1} - qq^\top)x_i = 0. \tag{40}$$

The quotient $\frac{\xi_i}{\sin\xi_i}$ makes sense in the interval $[0, \pi[$ by assuming that for $\xi_i = 0$, its value is equal to $\lim_{\xi_i \to 0^+} \frac{\xi_i}{\sin\xi_i} = 1$. Similar indeterminate forms that appear throughout the text will be treated the same way.

Remark 2. *From the equation that defines the Riemannian mean q, it is also clear that*

$$\Big(\sum_{i=0}^{2} \tfrac{\xi_i\cos\xi_i}{\sin\xi_i}\Big)q = \sum_{i=0}^{2} \tfrac{\xi_i}{\sin\xi_i}x_i, \tag{41}$$

showing that the Riemannian mean of the points x_0, x_1, x_2 belongs to the vector subspace of $\mathbb{R}^{n+1}$ spanned by them. In the particular case when the 3 points belong to the same geodesic in S^n, then q also belongs to the same geodesic.

There is one particular situation when the Riemannian mean is given explicitly in terms of the given points, as the following result shows.

Lemma 1 ([20])**.** *If x_0, x_1, x_2 are the vertices of a spherical equilateral triangle lying in $\mathcal{H}^n$, then the Riemannian mean of these points coincides with its spherical projected arithmetic mean, given explicitly by*

$$q = \frac{\sum_{i=0}^{2} x_i}{\left\|\sum_{i=0}^{2} x_i\right\|}. \tag{42}$$

Proof. Since the function Φ has a unique critical point, one just needs to show that the spherical projected arithmetic mean given by (42) satisfies the equation (40).

Let Θ denote the common length of the edges of the equilateral triangle, that is, $\Theta = \arccos\langle x_0, x_1\rangle = \arccos\langle x_0, x_2\rangle = \arccos\langle x_1, x_2\rangle$. So,

$$\|x_0 + x_1 + x_2\|^2 = 3\big(1 + 2\cos\Theta\big),$$

$$\langle x_0 + x_1 + x_2, x_i\rangle = 1 + 2\cos\Theta, \quad \text{for } i = 0, 1, 2,$$

and

$$\langle q, x_i\rangle = \sqrt{\tfrac{1+2\cos\Theta}{3}}, \quad \text{for } i = 0, 1, 2.$$

Consequently, $\frac{\xi_i}{\sin\xi_i}$ has the same value for $i = 0, 1, 2$, and the left hand side of equation (40) reduces to

$$\tfrac{\xi_0}{\sin\xi_0}(I - qq^\top)\Big(\sum_{i=0}^{2} x_i\Big). \tag{43}$$

Since $\sum_{i=0}^{2} x_i$ is a multiple of q, the expression in (43) is identically zero. $\square$

4.3. Defining the manifold M of all configurations

Consider the diffeomorphism that defines the Riemannian normal coordinates around a point q

$$\varphi_q \colon S^n \to T_q S^n, \quad p \mapsto \tfrac{\alpha}{\sin\alpha}(p - q\cos\alpha), \tag{44}$$

where $\cos\alpha = q^\top p$.

Let M be the smooth manifold consisting of all sets of pairs (p_1, p_2) such that (p_0, p_1, p_2) has the Riemannian mean q, that is

$$M = \Big\{(p_1, p_2) \in \mathcal{H}^n \times \mathcal{H}^n \;:\; (I - qq^\top)\sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} p_i = 0\Big\}, \tag{45}$$

where $\cos\alpha_i = q^\top p_i$ and I denotes the identity matrix of order $n+1$. In terms of Riemannian normal coordinates around q, M can be written as

$$M = \Big\{(p_1, p_2) \in \mathcal{H}^n \times \mathcal{H}^n \;:\; \sum_{i=0}^{2} \varphi_q(p_i) = 0\Big\}. \tag{46}$$

Theorem 3. *M is an $n-$dimensional smooth manifold.*

Proof. We will use the regular value theorem to show that M is the zero fiber of the following differentiable function

$$\begin{array}{rccl} \varphi: & \mathcal{H}^n \times \mathcal{H}^n & \longrightarrow & T_q\mathcal{H}^n \\ & (p_1, p_2) & \longmapsto & \varphi(p_1, p_2) = \sum_{i=0}^{2} \varphi_q(p_i). \end{array} \tag{47}$$

Given $(p_1, p_2) \in \varphi^{-1}(\{0\})$, the tangent map of φ at (p_1, p_2) is defined as

$$\begin{array}{rccl} D\varphi(p_1, p_2): & T_{p_1}\mathcal{H}^n \times T_{p_2}\mathcal{H}^n & \longrightarrow & T_q\mathcal{H}^n \\ & (v_1, v_2) & \longmapsto & \begin{bmatrix} \mathrm{D}\varphi_q(p_1) & \mathrm{D}\varphi_q(p_2) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \end{array} \tag{48}$$

Since $\mathrm{D}\varphi_q(p_i)$ is an isomorphism between $T_{p_i}\mathcal{H}^n$ and $T_q\mathcal{H}^n$, its rank is n and so $\mathrm{rank}(D\varphi(p_1, p_2)) = n$, showing that $D\varphi(p_1, p_2)$ is surjective everywhere. Consequently, φ is a submersion and $M = \varphi^{-1}(\{0\})$ is an n-dimensional smooth manifold. □

Visualizing the manifold M is challenging, even for dimension 2, but one can exhibit some symmetries.

Using the projection operator, M can be rewritten as

$$M = \big\{(p_1, p_2) \in \mathcal{H}^n \times \mathcal{H}^n \;:\; P_q^{\perp}\big(\sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} p_i\big) = 0\big\}. \tag{49}$$

Proposition 1. *Assume that $p_0 = q \neq p_1 \in S^n$. Then, $(p_1, R_q(p_1)) \in M$, where R_q is the reflection operator defined in* (37).

Proof. To show that $(p_1, p_2) \in M$ when $p_2 = R_q(p_1) := -p_1 + 2q\cos\alpha_1$, notice that in this case $\alpha_0 = 0$, and R_q leaves q invariant. This implies $\cos\alpha_2 = \cos\alpha_1$ and, consequently, $\frac{\alpha_2}{\sin\alpha_2} = \frac{\alpha_1}{\sin\alpha_1}$. So,

$$\sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} p_i = \big(q + \tfrac{\alpha_1}{\sin\alpha_1}(p_1 + R_q(p_1)\big) = \big(q + \tfrac{2\alpha_1\cos\alpha_1}{\sin\alpha_1} q\big),$$

showing that

$$(I - qq^{\top}) \sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} p_i = 0.$$

So, $(p_1, R_q(p_1)) \in M$, $\square$

Now, assume that $p_0 \neq q$, and let $R_{p_0 q}$ denote a reflection operator along the plane spanned by p_0, q and 0. With respect to this plane, any $y \in \mathbb{R}^{n+1}$ can be written as $y = y^{\perp} + \bar{y}$, where $\bar{y}$ is the orthogonal projection of y onto the plane. Then,

$$R_{p_0 q} \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \quad y \mapsto 2\bar{y} - y. \tag{50}$$

Moreover, $\bar{y} = \alpha p_0 + \beta q$, where

$$\alpha = \tfrac{1}{\sin^2\alpha_0}\big(p_0^{\top} y - (\cos\alpha_0) q^{\top} y\big), \quad \beta = \tfrac{1}{\sin^2\alpha_0}\big(q^{\top} y - (\cos\alpha_0) p_0^{\top} y\big),$$

or, equivalently, $\bar{y} = \frac{1}{\sin^2\alpha_0}\big(p_0 p_0^{\top} + qq^{\top} - \cos\alpha_0\,(p_0 q^{\top} + q p_0^{\top})\big)\, y$.

So, the orthogonal matrix A_R describing the reflection $R_{p_0 q}$ can be written as

$$A_R = \tfrac{2}{\sin^2\alpha_0}\big(p_0 p_0^{\top} + qq^{\top} - \cos\alpha_0\,(p_0 q^{\top} + q p_0^{\top})\big) - I_{n+1}.$$

Proposition 2. *If $p_0 \neq q$ and $(p_1, p_2) \in M$, then $(A_R(p_1), A_R(p_2)) \in M$, i.e., A_R is a symmetry of M.*

Proof. We first show that A_R commutes with $I - qq^\top$. This follows from the fact that $[p_0p_0^\top - \cos\alpha_0(p_0q^\top + qp_0^\top), qq^\top] = 0$, which is easily checked from

$$\begin{aligned}[p_0p_0^\top, qq^\top] &= \cos\alpha_0(p_0q^\top - qp_0^\top),\\ [p_0q^\top + qp_0^\top, qq^\top] &= p_0q^\top - qp_0^\top.\end{aligned}$$

Moreover, it is clear that A_R leaves p_0 and q invariant, and $q^\top A_R(p_i) = q^\top p_i = \cos\alpha_i$, for $i = 1, 2$. Consequently,

$$P_q^\perp\big(\sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} p_i\big) = 0 \quad \text{implies} \quad P_q^\perp\big(\sum_{i=0}^{2} \tfrac{\alpha_i}{\sin\alpha_i} A_R(p_i)\big) = 0,$$

that is, if $(p_1, p_2) \in M$ then $(A_R(p_1), A_R(p_2)) \in M$. □

4.4. Tangent and normal spaces to M

The following result will be useful.

Lemma 2. *Consider the function*

$$\begin{aligned} f : \mathcal{H}^n &\to \mathcal{H}^n \\ p_i &\mapsto \tfrac{\alpha_i}{\sin\alpha_i} p_i \end{aligned}. \tag{51}$$

The tangent map of f at the point p_i is given by

$$\begin{aligned} \mathrm{D}\, f(p_i) : T_{p_i}\mathcal{H}^n &\to T_{f(p_i)}\mathcal{H}^n \\ v_i &\mapsto \big(\tfrac{\alpha_i}{\sin\alpha_i} I + \big(\tfrac{\alpha_i\cos\alpha_i - \sin\alpha_i}{\sin^3\alpha_i}\big) p_i q^\top\big) v_i \end{aligned}. \tag{52}$$

Proof. The proof uses some elementary calculations based on the fact that for $\alpha_i = \arccos(q^\top p_i)$, $\mathrm{D}\alpha_i(p_i)(v_i) = \frac{-1}{\sin\alpha_i}(q^\top v_i)$. □

Now, it is convenient to define matrices

$$A_i = \tfrac{\alpha_i}{\sin\alpha_i} I + \big(\tfrac{\alpha_i\cos\alpha_i - \sin\alpha_i}{\sin^3\alpha_i}\big) p_i q^\top. \tag{53}$$

One can check that they are full rank. Indeed, following [21], p.475,

$$\det(A_i) = \big(\tfrac{\alpha_i}{\sin\alpha_i}\big)^{n+1}\big(\tfrac{\alpha_i - \cos\alpha_i\sin\alpha_i}{\alpha_i\sin^2\alpha_i}\big), \tag{54}$$

and since the second term in (54) is increasing in the interval $]0,\pi[$ and satisfies

$$\lim_{\alpha_i \to 0^+} \tfrac{\alpha_i - \cos\alpha_i \sin\alpha_i}{\alpha_i \sin^2\alpha_i} = 2/3,$$

we can conclude that $\det(A_i) \neq 0$, in the interval $[0,\pi[$. Its inverse can be derived from the Sherman-Morrison formula in [21], to obtain

$$A_i^{-1} = \tfrac{\sin\alpha_i}{\alpha_i}\big(I + \tfrac{\sin\alpha_i - \alpha_i\cos\alpha_i}{\alpha_i - \sin\alpha_i\cos\alpha_i} p_i q^\top\big).$$

Using the matrices A_i, the linear transformation $D\varphi(p_1,p_2)$ defined in (48) can be written as the following block matrix

$$D\varphi(p_1,p_2) = (I - qq^\top)\big[A_1\ A_2\big].$$

The tangent space to M at the point (p_1,p_2), can be characterized implicitly by

$$T_{(p_1,p_2)}M = \big\{(v_1,v_2) \in \mathbb{R}^{n+1}\times\mathbb{R}^{n+1} : X_1v_1 + X_2v_2 = 0,\ p_1^\top v_1 = 0,\ p_2^\top v_2 = 0\big\}, \tag{55}$$

where X_i, $i = 1,2$, is the rank-n matrix

$$X_i := (I - qq^\top)A_i. \tag{56}$$

Example 1. *If $(p_1,p_1) \in M$, the tangent space simplifies to*

$$T_{(p_1,p_1)}M = \big\{(v_1,-v_1)\ :\ v_1 \in T_{p_1}S^n\big\}.$$

Indeed, this vector space is n-dimensional and its vectors trivially satisfy the constraints in (55)*.*

Example 2. *If $p_0 = q \in \mathcal{H}^n$ and $p_1 \in \mathcal{H}^n$ satisfies $\cos^{-1}(q^\top p_1) \in]0,\pi/2[$, then*

$$T_{(p_1,R_q(p_1))}M = \big\{(v_1,R_q(v_1))\ :\ v_1 \in T_{p_1}S^n\big\},$$

where R_q is the reflection operator defined in (37)*.*

To show this, first notice that since $v_1^\top p_1 = 0$, also $(R_q(v_1))^\top R_q(p_1) = 0$, and so the last constraint in (55) *holds. To check that the first constraint in* (55) *also holds, compute X_1v_1 and X_2v_2, with $v_2 = R_q(v_1) = -v_1 + 2q^\top v_1 q$ and taking into consideration that if $p_2 := R_q(v_1)$, then $\cos\alpha_1 = \cos\alpha_2$, we conclude after simplifications that $X_2v_2 = -X_1v_1$, which completes the verification.*

In what follows, it is convenient to rewrite the tangent space in a more compact form. For that, define the the block matrix

$$X = \begin{bmatrix} X_1 & X_2 \\ p_1^\top & 0 \\ 0 & p_2^\top \end{bmatrix} \in \mathbb{R}^{(n+3)\times(2n+2)}.$$

Notice that the tangent space to M is the kernel of the matrix X, therefore the normal space is the row space (RS) of the matrix X (or the column space (CS) of $X^\top$), that is,

$$\begin{aligned} &T_{(p_1,p_2)}M = \ker(X) = \{x \in \mathbb{R}^{2n+2} \mid Xx = 0\ \}; \\ &T^\perp_{(p_1,p_2)}M = RS(X) = CS(X^\top) = \{X^\top x \mid\ x \in \mathbb{R}^{n+3}\}. \end{aligned} \tag{57}$$

By the rank-nullity theorem, we conclude that $\dim(T^\perp_{(p_1,p_2)}M) = n+2$.

Since M is an embedded submanifold of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$, it is convenient to define the projection operators onto the tangent and the normal spaces of M. For that, we first recall some facts about the Moore-Penrose inverse of X, denoted by X^+ (see, for instance, [11] or [21] for details).

The projection operators onto the tangent and the normal spaces of M are therefore defined, respectively, by

$$P^\perp_{\ker(X)} \colon \mathbb{R}^{n+1}\times\mathbb{R}^{n+1} \to T_{(p_1,p_2)}M, \quad (\omega_1,\omega_2) \mapsto (I_{2n+2} - X^+X)\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}, \tag{58}$$

$$P^\perp_{\mathrm{RS}(X)} \colon \mathbb{R}^{n+1}\times\mathbb{R}^{n+1} \to T^\perp_{(p_1,p_2)}M, \quad (\omega_1,\omega_2) \mapsto X^+X\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}. \tag{59}$$

4.5. Geodesics in M

In order to characterize the geodesics in M, let $\gamma : t \in (a,b) \mapsto \gamma(t) = (\gamma_1(t), \gamma_2(t)) \in M$. Then, for $i = 1, 2$,

$$\langle \gamma_i(t), \gamma_i(t)\rangle = 1,\ \forall t \in (a,b). \tag{60}$$

and

$$(I - qq^\top)\sum_{i=1}^{2} \tfrac{\alpha_i(t)}{\sin\alpha_i(t)}\gamma_i(t) + (I - qq^\top)\tfrac{\alpha_0}{\sin\alpha_0}p_0 = 0, \tag{61}$$

where $\alpha_i(t) = \arccos\langle q, \gamma_i(t)\rangle$, $\forall t \in (a, b)$, $i = 1, 2$.

Differentiating (60) and (61) with respect to t, one gets

$$\langle \gamma_i(t), \dot{\gamma}_i(t)\rangle = 0, \ \ i = 1, 2, \tag{62}$$

and

$$(I - qq^\top) \sum_{i=1}^{2} \big(\tfrac{\alpha_i(t)}{\sin \alpha_i(t)} I + \tfrac{\alpha_i(t) \cos \alpha_i(t) - \sin \alpha_i(t)}{\sin^3 \alpha_i(t)} \gamma_i(t) q^\top \big) \dot{\gamma}_i(t) = 0. \tag{63}$$

Introducing for $i = 1, 2$

$$X_i(t) = (I - qq^\top) \big(\tfrac{\alpha_i(t)}{\sin \alpha_i(t)} I + \tfrac{\alpha_i(t) \cos \alpha_i(t) - \sin \alpha_i(t)}{\sin^3 \alpha_i(t)} \gamma_i(t) q^\top \big),$$

and the block matrix

$$X(t) = \begin{bmatrix} X_1(t) & X_2(t) \\ \gamma_1(t)^\top & 0 \\ 0 & \gamma_2(t)^\top \end{bmatrix},$$

equations (62)-(63) are equivalent to

$$X(t)\dot{\gamma}(t) = 0, \quad \forall t \in (a, b). \tag{64}$$

If we differentiate (64) with respect to t, one also gets the following relation between the extrinsic acceleration (in the embedding space $\mathbb{R}^{2n+2}$) and the velocity of any curve in M

$$X(t)\ddot{\gamma}(t) = -\dot{X}(t)\dot{\gamma}(t). \tag{65}$$

In order for γ to be a geodesic in M, the orthogonal projection of the extrinsic acceleration $\ddot{\gamma}(t)$ onto the tangent space $T_{(\gamma_1(t), \gamma_2(t))}M$ should vanish, for all $t \in (a, b)$. So, using (58) one must have

$$\ddot{\gamma} - X^+ X \ddot{\gamma} = 0. \tag{66}$$

Using (65), the geodesic equation (66) can be rewritten as

$$\ddot{\gamma} + X^+ \dot{X} \dot{\gamma} = 0. \tag{67}$$

So far, no explicit solutions of this equation are known. Nevertheless we can state that geodesics on M are not pairs of geodesics on S^n, except possibly when all the points are aligned.

4.6. Cost function on M

Now, consider the smooth cost function

$$
\begin{aligned}
F\colon M &\longrightarrow \mathbb{R},\\
(p_1,p_2) &\longmapsto \tfrac{1}{4}\big(d^2(p_0,p_1)-d^2(p_0,p_2)\big)^2+\tfrac{1}{4}\big(d^2(p_0,p_1)-d^2(p_1,p_2)\big)^2\\
&\quad+\tfrac{1}{4}\big(d^2(p_0,p_2)-d^2(p_1,p_2)\big)^2\\
&\quad=\tfrac{1}{4}\big[\big(\Theta_1^2-\Theta_2^2\big)^2+\big(\Theta_1^2-\Theta_3^2\big)^2+\big(\Theta_2^2-\Theta_3^2\big)^2\big],
\end{aligned}
\tag{68}
$$

where $\Theta_i=\arccos\langle p_0,p_i\rangle$, $i=1,2$, and $\Theta_3=\arccos\langle p_1,p_2\rangle$.

Clearly if (p_0,p_1,p_2) describes a spherical equilateral triangle, then F attains its global minimum. Since the tangent space to M is only defined implicitly, we naturally extend F from M to a smooth function $\hat{F}$, then compute its derivative at the point (p_1,p_2) in the embedding space and use the implicit definition of the tangent space.

In order to simplify the expression for the differential of $\hat{F}$, introduce

$$
\begin{aligned}
\beta_1 &:= \Theta_2^2+\Theta_3^2-2\Theta_1^2\\
\beta_2 &:= \Theta_1^2+\Theta_3^2-2\Theta_2^2.
\end{aligned}
$$

Lemma 3.

$$
\begin{aligned}
\mathrm{D}\hat{F}(p_1,p_2)(v_1,v_2) &= \big(\tfrac{\Theta_1}{\sin\Theta_1}\beta_1 p_0^\top-\tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_2^\top\big)v_1\\
&\quad+\big(\tfrac{\Theta_2}{\sin\Theta_2}\beta_2 p_0^\top-\tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_1^\top\big)v_2.
\end{aligned}
\tag{69}
$$

Proof. In order to compute the tangent map of $\hat{F}$ at a pair (p_1,p_2), first notice that the expression of F is equivalent to

$$
F(p_1,p_2)=\tfrac{1}{2}\big(\Theta_1^4+\Theta_2^4+\Theta_3^4-\Theta_1^2\big(\Theta_2^2+\Theta_3^2\big)-\Theta_2^2\Theta_3^2\big).
\tag{70}
$$

Thus, given $(v_1,v_2)\in T_{(p_1,p_2)}M$, we can write

$$
\begin{aligned}
\mathrm{D}\hat{F}(p_1,p_2)(v_1,v_2) &= -2\tfrac{\Theta_1^3}{\sin\Theta_1}p_0^\top v_1-2\tfrac{\Theta_2^3}{\sin\Theta_2}p_0^\top v_2-2\tfrac{\Theta_3^3}{\sin\Theta_3}(p_2^\top v_1+p_1^\top v_2)\\
&\quad+\tfrac{\Theta_1}{\sin\Theta_1}\big(\Theta_2^2+\Theta_3^2\big)p_0^\top v_1+\tfrac{\Theta_1^2\Theta_2}{\sin\Theta_2}p_0^\top v_2+\tfrac{\Theta_3^2\Theta_2}{\sin\Theta_2}p_0^\top v_2\\
&\quad+\tfrac{\Theta_1^2\Theta_3}{\sin\Theta_3}\big(p_2^\top v_1+p_1^\top v_2\big)+\tfrac{\Theta_2^2\Theta_3}{\sin\Theta_3}\big(p_2^\top v_1+p_1^\top v_2\big)\\
&=\Big(\tfrac{\Theta_1}{\sin\Theta_1}\underbrace{\big(\Theta_2^2+\Theta_3^2-2\Theta_1^2\big)}_{:=\beta_1}p_0^\top+\tfrac{\Theta_3}{\sin\Theta_3}\underbrace{\big(\Theta_1^2+\Theta_2^2-2\Theta_3^2\big)}_{:=-(\beta_1+\beta_2)}p_2^\top\Big)v_1\\
&\quad+\Big(\tfrac{\Theta_2}{\sin\Theta_2}\underbrace{\big(\Theta_1^2+\Theta_3^2-2\Theta_2^2\big)}_{:=\beta_2}p_0^\top+\tfrac{\Theta_3}{\sin\Theta_3}\underbrace{\big(\Theta_1^2+\Theta_2^2-2\Theta_3^2\big)}_{:=-(\beta_1+\beta_2)}p_1^\top\Big)v_2
\end{aligned}
$$

□

The Euclidean gradient of $\mathrm{D}\hat{F}(p_1, p_2)$ is therefore given by

$$\nabla\widehat{F}(p_1,p_2) = \begin{bmatrix} \frac{\Theta_1}{\sin\Theta_1}\beta_1 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_2 \\ \frac{\Theta_2}{\sin\Theta_2}\beta_2 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_1 \end{bmatrix}, \tag{71}$$

and the Riemannian gradient of F is given by

$$\nabla F(p_1,p_2) = P^{\perp}_{\ker(X)} \begin{bmatrix} \frac{\Theta_1}{\sin\Theta_1}\beta_1 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_2 \\ \frac{\Theta_2}{\sin\Theta_2}\beta_2 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_1 \end{bmatrix}. \tag{72}$$

4.6.1. Critical points

Theorem 4. *In the three following situations, $(p_1, p_2) \in M$ is a critical point of the functional F defined by* (68)*.*

1. $p_0 = p_1 = p_2 = q$;
2. p_0, p_1, p_2 *form an equilateral spherical triangle;*
3. $p_1 = p_2$, *and moreover* $q = \frac{\sin(\frac{\Theta_1}{3})p_0 + \sin(\frac{2\Theta_1}{3})p_1}{\sin\Theta_1}$, *where* $\Theta_1 = \arccos(p_0^\top p_1)$.

Proof. Since $(p_1, p_2) \in M$ is a critical point of F if and only if the Riemannian gradient ∇F vanishes at that point, it is enough to show that in all these cases the orthogonal projection, onto the tangent space $T_{(p_1,p_2)}M$, of the Euclidean gradient $\nabla\widehat{F}$ (given in (71)) vanishes.

Case 1. $p_0 = p_1 = p_2 = q$.
In this case, $\Theta_i = \arccos(q^\top p_i) = \arccos(1) = 0$, $\forall i = 1, 2, 3$, and so $\beta_1 = \beta_2 = 0$. Consequently, $\nabla\widehat{F}(p_1, p_2) = 0$ and $\nabla F(p_1, p_2) = 0$.

Case 2. p_0, p_1, p_2 form an equilateral spherical triangle.
Here we have $\Theta_1 = \Theta_2 = \Theta_3$. So, as in the previous case, $\beta_1 = \beta_2 = 0$ and clearly also $\nabla\widehat{F}(p_1, p_2) = \nabla F(p_1, p_2) = 0$.

Case 3. $p_1 = p_2\,(\neq p_0)$ implies that $\alpha_1 = \alpha_2$, $\Theta_3 = 0$, $\Theta_1 = \Theta_2$, and according to Remark 2, all the points are on the same geodesic. Moreover, $\alpha_0 = 2\alpha_1$ and consequently $\cos\Theta_1 = \cos(3\alpha_1)$, $\Theta_1^2 = 9\alpha_1^2$, and $\beta_1 = \beta_2 = -9\alpha_1^2$.
We first obtain the value of q in terms of p_0 and p_1. For that, consider the geodesic in S^n going through q at $t = 0$, with velocity v, $\gamma(t) = \cos(t\|v\|)q + \frac{\sin(t\|v\|)}{\|v\|}v$. Assume that $p_0 = \gamma(t_0)$, $p_1 = \gamma(t_1)$, with $t_1 > 0$.
Then, $t_0 = -2t_1$ and, since $\cos\alpha_i = q^\top p_1$, we can write

$$p_0 = \cos(\alpha_0)q - \frac{\sin(\alpha_0)}{\|v\|}v, \quad p_1 = \cos(\alpha_1)q + \frac{\sin(\alpha_1)}{\|v\|}v.$$

Multiplying both sides of the first equation by $\sin\alpha_1$, both sides of the second equation by $\sin(2\alpha_1)$, adding them up, and using the fact that $\alpha_0 = 2\alpha_1$ and $\Theta_1 = 3\alpha_1$, one obtains

$$q = \frac{\sin(\frac{\Theta_1}{3})p_0 + \sin(\frac{2\Theta_1}{3})p_1}{\sin\Theta_1}.$$

In order to show that (p_1, p_1) is a critical point, it is enough to observe that in this case the Euclidean gradient in (71) reduces to

$$\nabla\widehat{F}(p_1, p_1) = \begin{bmatrix} \frac{\Theta_1}{\sin\Theta_1}\beta_1 p_0 - 2\frac{\Theta_3}{\sin\Theta_3}\beta_1 p_1 \\ \frac{\Theta_1}{\sin\Theta_1}\beta_1 p_0 - 2\frac{\Theta_3}{\sin\Theta_3}\beta_1 p_1 \end{bmatrix},$$

and clearly leaves in the orthogonal space to $T_{(p_1,p_1)}M$ given in Example 2. So, the Riemannian gradient vanishes at (p_1, p_1).

□

Theorem 5. *The only critical points of the functional F in M are those in the previous theorem.*

Proof. Based on the characterization of the normal space given in (57), we can also state that (p_1, p_2) is a critical point of F if, and only if, there exists $(x_1, x_2, x_3) \in \mathbb{R}^{n+1} \times \mathbb{R} \times \mathbb{R}$ such that

$$\begin{bmatrix} \frac{\Theta_1}{\sin\Theta_1}\beta_1 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_2 \\ \frac{\Theta_2}{\sin\Theta_2}\beta_2 p_0 - \frac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_1 \end{bmatrix} = \begin{bmatrix} X_1^\top & p_1 & 0 \\ X_2^\top & 0 & p_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \quad (73)$$

To prove this theorem we are going to show that in all the situations not covered by the previous theorem the system (73) is inconsistent or impossible. Let us assume that all the three points p_0, p_1 and p_2 are distinct from each other. Suppose, by contradiction, that the system (73) is possible. This means that $\nabla\widehat{F}(p_1, p_2) \in RS(X)$, which is equivalent to have $P^\perp_{\ker(X)}(\nabla\widehat{F}(p_1, p_2)) = 0$, where $P^\perp_{\ker(X)}$ is the orthogonal projection operator defined by (58). Therefore, $\nabla\widehat{F}(p_1, p_2)$ is a solution of the homogeneous system

$$(I - X^+X)y = 0. \quad (74)$$

Now, let $R = \mathrm{rref}(I - X^+X)$ be the reduced row echelon form of $I - X^+X$ (see, for instance, [11] for details). Then, there exists an (invertible) product

$\mathcal{E}$ of the elementary row operations such that $R = \mathcal{E}(I - X^+X)$. So, it is immediate to see that y is a solution of (74) if and only if it is a solution of $Ry = 0$. But, since $\operatorname{rank}(I - X^+X) = n$, then R has also rank n and can be represented as

$$R = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, \tag{75}$$

where the last $n+2$ rows are null. Therefore, in order for

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \tag{76}$$

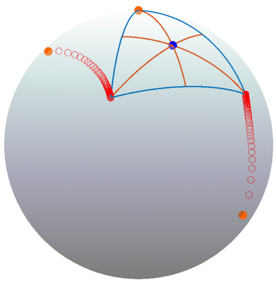
where $y_1 \in \mathbb{R}^n$ and $y_2 \in \mathbb{R}^{n+2}$, to be a solution for $Ry = 0$ it is necessary that $y_2 = 0$. It is evident that since $p_i \in S^n$, $i = 0, 1, 2$, $\nabla\widehat{F}(p_1, p_2)$ cannot fulfill this requirement, which leads us to the desired contradiction. □

Next, we present the steepest descent algorithm (Algorithm 3) in order to obtain approximate solutions to the problem. Figure 4 illustrates this method for points (p_1, p_2) in M with fixed $p_0 = (0, 0, 1)$. In the first two images, the points (whose coordinates are given in each caption) converge to the vertices of a spherical equilateral triangle, corresponding to the critical point stated in Theorem 4, case 2. In the third image, the points belong to the geodesic that contains p_0 and q and converge to the critical point, corresponding to case 3 in Theorem 4. The graphs on the right hand side show how the Euclidean distances between pairs of points (p_1, p_2) in successive iterations evolve. As expected, these distances converge linearly to zero.

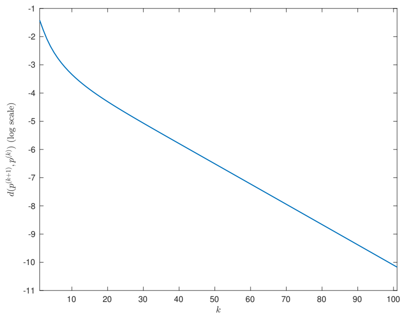
Algorithm 3: Steepest descent with Armijo line search

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)})$ and tolerance tol
Output: Stationary point $p^* = (p_1^*, p_2^*)$

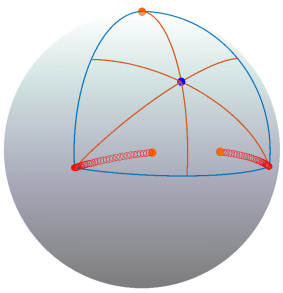
for $k = 0, 1, \ldots$ **do**
  Set $v_{(k)} = -\nabla F(p^{(k)})$;
  Determine the step length α_k according to Armijo rule;
  Set $p^{(k+1)} = \gamma(\alpha_k)$, where γ is a geodesic in M starting in $p^{(k)}$ with velocity $v_{(k)}$;
  Stop if $F(p^{(k)}) <$ tol or $\|\nabla F(p^{(k)})\| <$ tol
end



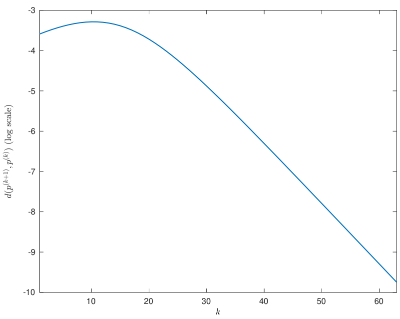
(a) $p_1 = (0, -\frac{\sqrt{3}}{2}, -\frac{1}{2})$, $p_2 = (-\frac{1}{2}, \frac{1}{2}, \frac{\sqrt{2}}{2})$



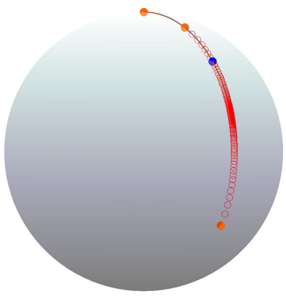
(b) Iteration k versus Euclidean distance



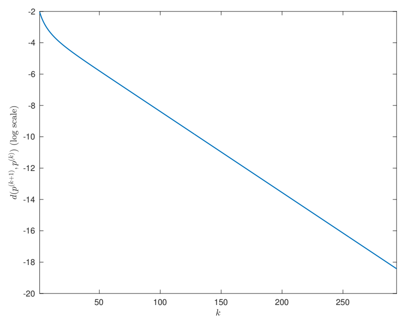
(c) $p_1 = (0, 1, 0)$, $p_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$



(d) Iteration k versus Euclidean distance



(e) $p_1 = (\frac{1}{\sqrt{2}} \sin \frac{\sqrt{2}}{3}, \frac{1}{\sqrt{2}} \sin \frac{\sqrt{2}}{3}, \cos \frac{\sqrt{2}}{3})$, $p_2 = (\frac{1}{\sqrt{2}} \sin \frac{3\sqrt{2}}{2}, \frac{1}{\sqrt{2}} \sin \frac{3\sqrt{2}}{2}, \cos \frac{3\sqrt{2}}{2})$



(f) Iteration k versus Euclidean distance

Figure 4: Plots obtained using Algorithm 3

4.6.2. Riemannian Hessian

According to the definition (see, for instance, p. 109 in [1] or Proposition 16.22. in [9]), the Riemannian Hessian along geodesics can be computed as

$$H_F(p_1, p_2)(v_1, v_2) = \tfrac{d^2}{dt^2}\Big|_{t=0} F(\gamma(t)), \tag{77}$$

where $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ is a geodesic in M satisfying $\gamma(0) = (p_1, p_2)$ and $\dot{\gamma}(0) = (v_1, v_2)$.

The presence of the Moore-Penrose inverse of X in the projection operator (58) makes the computation of the Riemannian Hessian of F challenging. Formula (7) in [2] provides an alternative to obtain this Hessian. However, in our case, the non-differentiability of $X^+(t)$ that appears in the projection operator (58) becomes an obstacle to use that formula.

In order to use (77), we need some additional computations to obtain $\dot{X}_i(0)$ and $\ddot{\Theta}_j(0)$.

Using equation (67), one can write

$$\begin{bmatrix} \ddot{\gamma}_1(0) \\ \ddot{\gamma}_2(0) \end{bmatrix} = -X^+(0)\dot{X}(0) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \tag{78}$$

where

$$X(t) = \begin{bmatrix} X_1(t) & X_2(t) \\ \gamma_1(t)^\top & 0 \\ 0 & \gamma_2(t)^\top \end{bmatrix},$$

and, for $i = 1, 2$, $X_i(t)$ is defined as before by

$$X_i(t) = (I - qq^\top)\big(\tfrac{\alpha_i(t)}{\sin \alpha_i(t)} I + \tfrac{\alpha_i(t)\cos\alpha_i(t) - \sin\alpha_i(t)}{\sin^3 \alpha_i(t)} \gamma_i(t) q^\top\big),$$

and

$$\alpha_i(t) = \arccos\langle q, \gamma_i(t)\rangle.$$

The derivatives evaluated at $t = 0$ of the two above expressions give

$$\dot{\alpha}_i(0) = -\tfrac{\langle q, v_i\rangle}{\sin \alpha_i},$$

and

$$
\begin{aligned}
\dot{X}_i(0) = &-\tfrac{(I-qq^\top)}{\sin^3\alpha_i}\big((q^\top v_i)(\sin\alpha_i - \alpha_i\cos\alpha_i)I + (\sin\alpha_i - \alpha_i\cos\alpha_i)v_i q^\top \\
&+ \tfrac{3\cos\alpha_i\sin\alpha_i - 3\alpha_i\cos^2\alpha_i - \alpha_i\sin^2\alpha_i}{\sin^2\alpha_i}(q^\top v_i)p_i q^\top\big) \\
= &-\tfrac{(I-qq^\top)}{\sin^3\alpha_i}\big((\sin\alpha_i - \alpha_i\cos\alpha_i)((q^\top v_i)I + v_i q^\top) \\
&+ (3\cot\alpha_i(1-\alpha_i\cot\alpha_i) - \alpha_i)(q^\top v_i)p_i q^\top\big)
\end{aligned}
\tag{79}
$$

Let $\Theta_i(t) = \arccos\langle p_0, \gamma_i(t)\rangle,\ i = 1, 2$ and $\Theta_3(t) = \arccos\langle \gamma_1(t), \gamma_2(t)\rangle$. Then, for $i = 1, 2$,

$$
-\dot{\Theta}_i(t)\sin\Theta_i(t) = \langle p_0, \dot{\gamma}_i(t)\rangle, \tag{80}
$$

and

$$
-\ddot{\Theta}_i(t)\sin\Theta_i(t) - \dot{\Theta}_i(t)^2\cos\Theta_i(t) = \langle p_0, \ddot{\gamma}_i(t)\rangle. \tag{81}
$$

So, evaluating the above at $t = 0$, yields

$$
\dot{\Theta}_i(0) = -\tfrac{\langle p_0, v_i\rangle}{\sin\Theta_i}, \quad \ddot{\Theta}_i(0) = -\tfrac{\cos\Theta_i\langle p_0, v_i\rangle^2 + \sin^2\Theta_i\langle p_0, \ddot{\gamma}_i(0)\rangle}{\sin^3\Theta_i}. \tag{82}
$$

Analogous computations for Θ_3 show that

$$
\dot{\Theta}_3(0) = -\tfrac{\langle p_2, v_1\rangle + \langle p_1, v_2\rangle}{\sin\Theta_3}, \tag{83}
$$

and

$$
\begin{aligned}
\ddot{\Theta}_3(0) &= -\tfrac{1}{\sin\Theta_3}\big(\dot{\Theta}_3^2(0)\cos\Theta_3 + \langle\ddot{\gamma}_1(0), p_2\rangle + 2\langle v_1, v_2\rangle + \langle\ddot{\gamma}_2(0), p_1\rangle\big) \\
&= -\tfrac{1}{\sin^3\Theta_3}\big(\cos\Theta_3\big(\langle p_1, v_2\rangle^2 + 2\langle p_1, v_2\rangle\langle p_2, v_1\rangle + \langle p_2, v_1\rangle^2\big) \\
&\qquad + \sin^2\Theta_3(\langle\ddot{\gamma}_1(0), p_2\rangle + 2\langle v_1, v_2\rangle + \langle\ddot{\gamma}_2(0), p_1\rangle)\big)
\end{aligned}
\tag{84}
$$

Then

$$
\begin{aligned}
&H_F(p_1, p_2)(v_1, v_2) = \tfrac{d^2}{dt^2}\big|_{t=0} F(\gamma(t)) \\
&= \tfrac{d}{dt}\big|_{t=0}\big(2\dot{\Theta}_1\Theta_1^3 + 2\dot{\Theta}_2\Theta_2^3 + 2\dot{\Theta}_3\Theta_3^3 - \dot{\Theta}_1\Theta_1\big(\Theta_2^2 + \Theta_3^2\big) - \dot{\Theta}_2\Theta_2\big(\Theta_1^2 + \Theta_3^2\big) \\
&\qquad - \dot{\Theta}_3\Theta_3\big(\Theta_1^2 + \Theta_2^2\big)\big) \\
&= \ddot{\Theta}_1(0)\Theta_1(2\Theta_1^2 - \Theta_2^2 - \Theta_3^2) + \ddot{\Theta}_2(0)\Theta_2(2\Theta_2^2 - \Theta_1^2 - \Theta_3^2) \\
&\quad + \ddot{\Theta}_3(0)\Theta_3(2\Theta_3^2 - \Theta_1^2 - \Theta_2^2) + \dot{\Theta}_1(0)^2\big(6\Theta_1^2 - \Theta_2^2 - \Theta_3^2\big) \\
&\quad + \dot{\Theta}_2(0)^2\big(6\Theta_2^2 - \Theta_1^2 - \Theta_3^2\big) + \dot{\Theta}_3(0)^2\big(6\Theta_3^2 - \Theta_1^2 - \Theta_2^2\big) \\
&\quad - 4\Theta_1\Theta_2\dot{\Theta}_1(0)\dot{\Theta}_2(0) - 4\Theta_1\Theta_3\dot{\Theta}_1(0)\dot{\Theta}_3(0) - 4\Theta_2\Theta_3\dot{\Theta}_2(0)\dot{\Theta}_3(0)
\end{aligned}
\tag{85}
$$

Now, plugging the expressions (82) and (84) in the latter and after rearranging the terms, one gets

$$\begin{aligned}
&H_F(p_1,p_2)(v_1,v_2) = \\
&= v_1^\top\big(\tfrac{\Theta_1\cos\Theta_1\beta_1+\sin\Theta_1(6\Theta_1^2-\Theta_2^2-\Theta_3^2)}{\sin^3\Theta_1}p_0p_0^\top \\
&+ \tfrac{\sin\Theta_3(6\Theta_3^2-\Theta_1^2-\Theta_3^2)-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_2p_2^\top - 2\tfrac{\Theta_1}{\sin\Theta_1}\tfrac{\Theta_3}{\sin\Theta_3}(p_0p_2^\top + p_2p_0^\top)\big)v_1 \\
&+ v_2^\top\big(\tfrac{\Theta_2\cos\Theta_2\beta_2+\sin\Theta_2(6\Theta_2^2-\Theta_1^2-\Theta_3^2)}{\sin^3\Theta_2}p_0p_0^\top \\
&+ \tfrac{\sin\Theta_3(6\Theta_3^2-\Theta_1^2-\Theta_2^2)-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_1p_1^\top - 2\tfrac{\Theta_2}{\sin\Theta_2}\tfrac{\Theta_3}{\sin\Theta_3}(p_0p_1^\top + p_1p_0^\top)\big)v_2 \\
&+ 2v_1^\top\big(\tfrac{(6\Theta_3^2-\Theta_1^2-\Theta_3^2)\sin\Theta_3-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_2p_1^\top - (\beta_1+\beta_2)\tfrac{\Theta_3}{\sin\Theta_3}I \\
&- 2\tfrac{\Theta_1}{\sin\Theta_1}\tfrac{\Theta_2}{\sin\Theta_2}p_0p_0^\top - 2\tfrac{\Theta_3}{\sin\Theta_3}\big(\tfrac{\Theta_1}{\sin\Theta_1}p_0p_1^\top + \tfrac{\Theta_2}{\sin\Theta_2}p_2p_0^\top\big)\big)v_2 + \tfrac{\Theta_1}{\sin\Theta_1}\beta_1\langle p_0,\ddot\gamma_1(0)\rangle \\
&+ \tfrac{\Theta_2}{\sin\Theta_2}\beta_2\langle p_0,\ddot\gamma_2(0)\rangle - \tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)(\langle p_2,\ddot\gamma_1(0)\rangle + \langle p_1,\ddot\gamma_2(0)\rangle)
\end{aligned} \tag{86}$$

So, if we denote by

$$\begin{aligned}
Y &= \tfrac{\Theta_1\cos\Theta_1\beta_1+\sin\Theta_1(6\Theta_1^2-\Theta_2^2-\Theta_3^2)}{\sin^3\Theta_1}p_0p_0^\top \\
&\quad+ \tfrac{(6\Theta_3^2-\Theta_1^2-\Theta_2^2)\sin\Theta_3-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_2p_2^\top - 2\tfrac{\Theta_1}{\sin\Theta_1}\tfrac{\Theta_3}{\sin\Theta_3}(p_0p_2^\top + p_2p_0^\top) \\
Z &= \tfrac{(6\Theta_3^2-\Theta_1^2-\Theta_2^2)\sin\Theta_3-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_2p_1^\top - (\beta_1+\beta_2)\tfrac{\Theta_3}{\sin\Theta_3}I \\
&\quad- 2\tfrac{\Theta_1}{\sin\Theta_1}\tfrac{\Theta_2}{\sin\Theta_2}p_0p_0^\top - 2\tfrac{\Theta_3}{\sin\Theta_3}\big(\tfrac{\Theta_1}{\sin\Theta_1}p_0p_1^\top + \tfrac{\Theta_2}{\sin\Theta_2}p_2p_0^\top\big) \\
W &= \tfrac{\Theta_2\cos\Theta_2\beta_2+\sin\Theta_2(6\Theta_2^2-\Theta_1^2-\Theta_3^2)}{\sin^3\Theta_2}p_0p_0^\top \\
&\quad+ \tfrac{\sin\Theta_3(6\Theta_3^2-\Theta_1^2-\Theta_2^2)-\Theta_3\cos\Theta_3(\beta_1+\beta_2)}{\sin^3\Theta_3}p_1p_1^\top - 2\tfrac{\Theta_2}{\sin\Theta_2}\tfrac{\Theta_3}{\sin\Theta_3}(p_0p_1^\top + p_1p_0^\top),
\end{aligned} \tag{87}$$

the Hessian of F at (p_1, p_2) can be written as

$$\begin{aligned}
&H_F(p_1,p_2)(v_1,v_2) = \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix}\begin{bmatrix} Y & Z \\ Z^\top & W \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\
&+ \tfrac{\Theta_1}{\sin\Theta_1}\beta_1 p_0^\top\ddot\gamma_1(0) + \tfrac{\Theta_2}{\sin\Theta_2}\beta_2 p_0^\top\ddot\gamma_2(0) - \tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)(p_2^\top\ddot\gamma_1(0) + p_1^\top\ddot\gamma_2(0)) \\
&= \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix}\begin{bmatrix} Y & Z \\ Z^\top & W \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\
&+ \begin{bmatrix} \tfrac{\Theta_1}{\sin\Theta_1}\beta_1 p_0^\top - \tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_2^\top & \tfrac{\Theta_2}{\sin\Theta_2}\beta_2 p_0^\top - \tfrac{\Theta_3}{\sin\Theta_3}(\beta_1+\beta_2)p_1^\top \end{bmatrix}\begin{bmatrix} \ddot\gamma_1(0) \\ \ddot\gamma_2(0) \end{bmatrix}.
\end{aligned} \tag{88}$$

Now, using equation (78) and taking into account the expression for the Euclidean gradient of $\widehat{F}$ given in (71), we can write the Hessian in terms of the coordinates of the embedding space

$$\begin{aligned} H_F(p_1,p_2)(v_1,v_2) &= \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix} \begin{bmatrix} Y & Z \\ Z^\top & W \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &- \nabla\widehat{F}(p_1,p_2)^\top \begin{bmatrix} X_1 & X_2 \\ p_1^\top & 0 \\ 0 & p_2^\top \end{bmatrix}^+ \begin{bmatrix} \dot{X}_1(0) & \dot{X}_2(0) \\ v_1^\top & 0 \\ 0 & v_2^\top \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \end{aligned} \tag{89}$$

which, according to the expressions for $\dot{X}_i(0)$ given by (79), is clearly a quadratic form in v_1 and v_2, although not in the canonical form.

Algorithm 5: Quasi-Newton method

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)})$, $\lambda > 0$ and tolerance tol

Output: Stationary point $p^* = (p_1^*, p_2^*)$

for $k = 0, 1, \ldots$ **do**

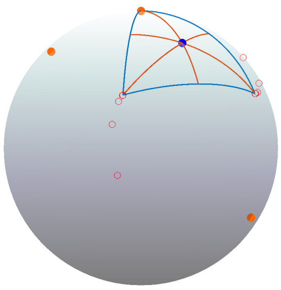
  Set $B^{(k)} = H_F(p^{(k)}) + \lambda I_n$;

  Solve $d^{(k)}$ from $B^{(k)} d^{(k)} = -\nabla F(p^{(k)})$;

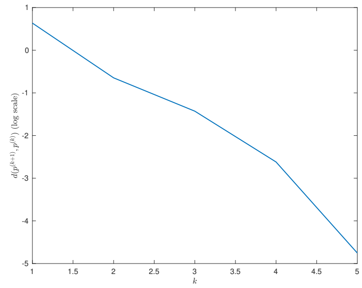
  Update $p^{(k+1)} = \gamma(1)$, where γ is a geodesic in M starting in $p^{(k)}$ with velocity $d^{(k)}$;

  Stop if $F(p^{(k)}) < \mathsf{tol}$ or $\|\nabla F(p^{(k)})\| < \mathsf{tol}$

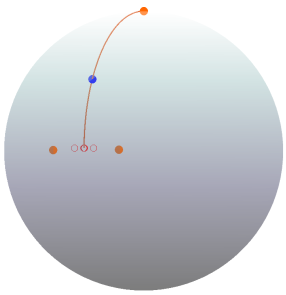
end



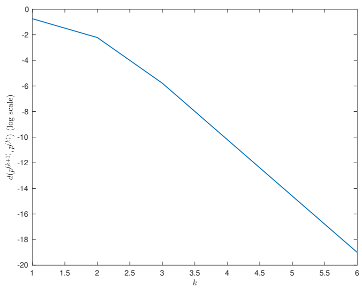
(a) $p_1 = (0, -\frac{\sqrt{3}}{2}, -\frac{1}{2})$, $p_2 = (-\frac{1}{2}, \frac{1}{2}, \frac{\sqrt{2}}{2})$



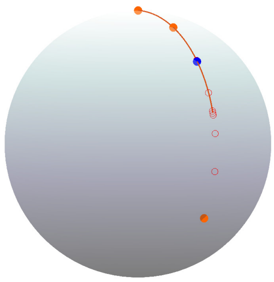
(b) Iteration k versus Euclidean distance



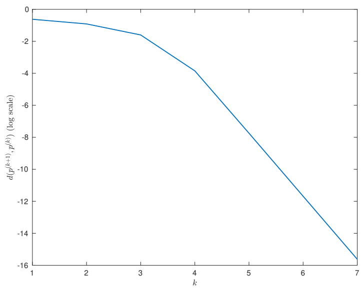
(c) $p_1 = (0, 1, 0)$, $p_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$



(d) Iteration k versus Euclidean distance



(e) $p_1 = (\frac{1}{\sqrt{2}} \sin \frac{\sqrt{2}}{3}, \frac{1}{\sqrt{2}} \sin \frac{\sqrt{2}}{3}, \cos \frac{\sqrt{2}}{3})$, $p_2 = (\frac{1}{\sqrt{2}} \sin \frac{3\sqrt{2}}{2}, \frac{1}{\sqrt{2}} \sin \frac{3\sqrt{2}}{2}, \cos \frac{3\sqrt{2}}{2})$



(f) Iteration k versus Euclidean distance

Figure 5: Plots obtained using Algorithm 5

We are now in conditions to classify the critical points of F given in Theorem 4.

Theorem 6. *The critical points of the function F on M are classified in the following way:*

1. *when $p_1 = p_2 \neq p_0$, the critical point (p_1, p_1) is a saddle point;*
2. *when p_0, p_1, p_2 form the vertices of a spherical triangle, the critical point (p_1, p_2) is a global minimum.*

Proof. For $p_1 = p_2 \neq p_0$, we have $\Theta_1 = \Theta_2 \neq 0$, $\Theta_3 = 0$, $\beta_1 = \beta_2 = -\Theta_1^2$. So, taking into consideration that

$$\lim_{\Theta_3 \to 0} \tfrac{\sin\Theta_3 - \Theta_3\cos\Theta_3}{\sin^3\Theta_3} = \tfrac{1}{3} \quad \text{and} \quad \lim_{\Theta_3 \to 0} \tfrac{\Theta_3}{\sin\Theta_3} = 1,$$

the expressions appearing in (87) reduce to

$$\begin{aligned} Y &= \tfrac{5\Theta_1^2\sin\Theta_1 - \Theta_1^3\cos\Theta_1}{\sin^3\Theta_1} p_0p_0^\top + \left(6 - \tfrac{2}{3}\Theta_1^2\right)p_1p_1^\top - \tfrac{2\Theta_1}{\sin\Theta_1}(p_0p_1^\top + p_1p_0^\top), \\ Z &= \left(6 - \tfrac{2}{3}\Theta_1^2\right)p_1p_1^\top - 2\tfrac{\Theta_1^2}{\sin^2\Theta_1}p_0p_0^\top - \tfrac{2\Theta_1}{\sin\Theta_1}(p_0p_1^\top + p_1p_0^\top) + 2\Theta_1^2 I_{n+1}, \\ W &= Y. \end{aligned} \tag{90}$$

To simplify notations, for the rest of the proof of statement 1., we will use Θ and α instead of Θ_1 and α_1, respectively. Although for this case $\alpha = \frac{\Theta}{3}$, in the calculations below we may use either Θ or α depending on which makes the expressions look simpler.

To show that (p_1, p_1) is a saddle point, it is enough to choose two different directions $(v_1, -v_1)$ in $T_{(p_1,p_1)}M$ for which the quadratic form

$$\begin{aligned} H_F(p_1,p_1)(v_1,-v_1) &= \begin{bmatrix} v_1^\top & -v_1^\top \end{bmatrix} \begin{bmatrix} Y & Z \\ Z & Y \end{bmatrix} \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix} \\ &\quad - \nabla\widehat{F}(p_1,p_1)^\top \begin{bmatrix} X_1 & X_1 \\ p_1^\top & 0 \\ 0 & p_1^\top \end{bmatrix}^+ \begin{bmatrix} \dot{X}_1(0) & \dot{X}_1(0) \\ v_1^\top & 0 \\ 0 & -v_1^\top \end{bmatrix} \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix} \\ &= 2v_1^\top(Y - Z)v_1 \\ &\quad - \begin{bmatrix} 2\Theta^2 p_1^\top - \tfrac{\Theta^3}{\sin\Theta}p_0^\top & 2\Theta^2 p_1^\top - \tfrac{\Theta^3}{\sin\Theta}p_0^\top \end{bmatrix} \begin{bmatrix} X_1 & X_1 \\ p_1^\top & 0 \\ 0 & p_1^\top \end{bmatrix}^+ \begin{bmatrix} 0 \\ v_1^\top v_1 \\ v_1^\top v_1 \end{bmatrix} \end{aligned} \tag{91}$$

has opposite signs.

According to the expressions for Y and Z, we conclude after simplifications that

$$Y - Z = \left(\tfrac{7\Theta^2 \sin\Theta - \Theta^3 \cos\Theta}{\sin^3\Theta}\right) p_0 p_0^\top - 2\Theta^2 I_{n+1}. \tag{92}$$

In order to evaluate the second expression in (91), let us introduce the matrix

$$W_1 = 2X_1^\top X_1 + p_1 p_1^\top. \tag{93}$$

According to the definition of X_1 given in (56), W_1 can be written as

$$\begin{aligned} W_1 &= \tfrac{2\alpha^2}{\sin^2\alpha} I + \tfrac{2\alpha(\alpha\cos\alpha - \sin\alpha)}{\sin^4\alpha}(p_1 q^\top + q p_1^\top) + \tfrac{2\sin^2\alpha - 2\alpha^2}{\sin^4\alpha} q q^\top + p_1 p_1^\top \\ &= \tfrac{2\alpha^2}{\sin^2\alpha}\left(I + \tfrac{\alpha\cos\alpha - \sin\alpha}{\alpha\sin^2\alpha}(p_1 q^\top + q p_1^\top) + \tfrac{\sin^2\alpha - \alpha^2}{\alpha^2\sin^2\alpha} q q^\top + \tfrac{\sin^2\alpha}{2\alpha^2} p_1 p_1^\top\right). \end{aligned} \tag{94}$$

By defining the following rank-two matrix

$$\begin{aligned} K :&= \tfrac{\alpha\cos\alpha - \sin\alpha}{\alpha\sin^2\alpha}(p_1 q^\top + q p_1^\top) + \tfrac{\sin^2\alpha - \alpha^2}{\alpha^2\sin^2\alpha} q q^\top + \tfrac{\sin^2\alpha}{2\alpha^2} p_1 p_1^\top \\ &= \begin{bmatrix} p_1 & q \end{bmatrix} \begin{bmatrix} \frac{\sin^2\alpha}{2\alpha^2} & \frac{\alpha\cos\alpha - \sin\alpha}{\alpha\sin^2\alpha} \\ \frac{\alpha\cos\alpha - \sin\alpha}{\alpha\sin^2\alpha} & \frac{\sin^2\alpha - \alpha^2}{\alpha^2\sin^2\alpha} \end{bmatrix} \begin{bmatrix} p_1^\top \\ q^\top \end{bmatrix}, \end{aligned} \tag{95}$$

we get

$$W_1 = \tfrac{2\alpha^2}{\sin^2\alpha}(I + K). \tag{96}$$

It turns out that W_1 is nonsingular and its inverse can be computed in closed form using Sherman-Morrison's formula recursively, as explained in [22]. Eventually, we obtain

$$W_1^{-1} = \tfrac{\sin^2\alpha}{2\alpha^2}\left(I - \tfrac{1}{d+e}(dK - K^2)\right), \tag{97}$$

with $d = 1 + \mathrm{tr}K$ and $e = \left((\mathrm{tr}K)^2 - \mathrm{tr}(K^2)\right)/2$.

Further computations lead to

$$\begin{aligned} d &= 1 + \tfrac{1}{2\alpha^2}\left(\sin^2\alpha + 4\alpha\cot\alpha(\alpha\cot\alpha - 1)\right), \\ d + e &= \tfrac{\sin^4\alpha}{2\alpha^4}. \end{aligned} \tag{98}$$

By using the well known properties of the Moore-Penrose inverse, it can be shown that

$$\begin{bmatrix} X_1 & X_1 \\ p_1^\top & 0 \\ 0 & p_1^\top \end{bmatrix}^+ = \frac{1}{2}\begin{bmatrix} 2W_1^{-1}X_1^\top & W_1^{-1}p_1 + p_1 & W_1^{-1}p_1 - p_1 \\ 2W_1^{-1}X_1^\top & W_1^{-1}p_1 - p_1 & W_1^{-1}p_1 + p_1 \end{bmatrix}. \tag{99}$$

We proceed by choosing

$$v_1 = \tfrac{\Theta}{\sin\Theta}\big(p_0 - p_1\cos\Theta\big),$$

i.e., v_1 is the initial velocity vector of the geodesic in S^n joining p_1 (at $t=0$) to p_0 (at $t=1$). Replacing this in (92), the Hessian given in (91) becomes

$$\begin{aligned} &H_F(p_1,p_1)(v_1,-v_1) = \\ &= 2\Theta^4\big(5 - \tfrac{\Theta}{\sin\Theta}\cos\Theta\big) - 2\Theta^4\big(2p_1^\top - \tfrac{\Theta}{\sin\Theta}p_0^\top\big)W_1^{-1}p_1 \\ &= 2\Theta^4\big(5 - \tfrac{\Theta}{\sin\Theta}\cos\Theta - 2p_1^\top W_1^{-1}p_1 + \tfrac{\Theta}{\sin\Theta}p_0^\top W_1^{-1}p_1\big). \end{aligned} \tag{100}$$

Using the expression for W_1^{-1} given by (97) and some standard trigonometric identities, it follows that

$$\begin{aligned} &p_1^\top W_1^{-1}p_1 = 1, \\ &p_0^\top W_1^{-1}p_1 = -2\cos\alpha + \tfrac{3\alpha}{\sin\alpha} - 4\alpha\sin\alpha. \end{aligned} \tag{101}$$

Plugging this expression into (100) and simplifying, one gets

$$H_F(p_1,p_1)(v_1,-v_1) = \tfrac{3^5\alpha^4}{2}\big(4\alpha^2 + \big(1-2\alpha\cot\alpha\big)^2 + 3\big), \tag{102}$$

which is greater than zero.

On the other hand, choosing $v_1 \in T_{p_1}S^n$ to be orthogonal to both p_1 and p_0, the quadratic form (91) simplifies to

$$\begin{aligned} H_F(p_1,p_1)(v_1,-v_1) &= -\|v_1\|^2\big(4\Theta^2 + 2\Theta^2 p_1^\top W_1^{-1}p_1 - \tfrac{\Theta^3}{\sin\Theta}p_0^\top W_1^{-1}p_1\big) \\ &= -\|v_1\|^2\Theta^2\big(6 - \tfrac{3\alpha}{\sin(3\alpha)}\tfrac{\cos(3\alpha)-\cos\alpha+2\,\alpha\sin(3\alpha)}{2\sin^2\alpha}\big), \\ &= -3^3\alpha^2\|v_1\|^2\big(2 - \tfrac{\alpha^2}{\sin^2\alpha} + \tfrac{2}{3}\tfrac{3\alpha}{\sin(3\alpha)}\cos\alpha\big), \end{aligned} \tag{103}$$

which is negative because $\frac{\alpha^2}{\sin^2\alpha} < 2$, for all $\alpha \in (0, \frac{\pi}{3})$.
This proves that (p_1,p_1) is indeed a saddle point.

We now proceed to the second case, when $\Theta_1 = \Theta_2 = \Theta_3$, and so $\beta_1 = \beta_2 = 0$. In this case, $\nabla\widehat{F}(p_1,p_2) = 0$ and the Hessian given by (89) reduces to the first term. In this case, formulas (87) simplify to

$$\begin{aligned} Y &= \tfrac{2\Theta_1^2}{\sin^2\Theta_1}\big(2p_0p_0^\top + 2p_2p_2^\top - p_0p_2^\top - p_2p_0^\top\big), \\ Z &= \tfrac{2\Theta_1^2}{\sin^2\Theta_1}\big(2p_2p_1^\top - p_0p_0^\top - p_0p_1^\top - p_2p_0^\top\big), \\ W &= \tfrac{2\Theta_1^2}{\sin^2\Theta_1}\big(2p_0p_0^\top + 2p_1p_1^\top - p_0p_1^\top - p_1p_0^\top\big). \end{aligned} \tag{104}$$

Let $(v_1, v_2) \in T_{(p_1,p_2)}M$. Then, after computations and simplifications, we can write

$$\begin{aligned} H_F(p_1,p_2)(v_1,v_2) &= \begin{bmatrix} v_1^\top & v_2^\top \end{bmatrix} \begin{bmatrix} Y & Z \\ Z^\top & W \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \\ &= \tfrac{2\Theta_1^2}{\sin^2\Theta_1} \big((v_1^\top p_0 - v_1^\top p_2)^2 + (v_2^\top p_0 - v_2^\top p_1)^2 + (v_1^\top p_0 - v_2^\top p_1)^2 \\ &\quad + (v_1^\top p_2 - v_2^\top p_0)^2 + 4(v_1^\top p_2)(v_2^\top p_1) - 2(v_1^\top p_0)(v_2^\top p_0) \big). \end{aligned} \tag{105}$$

To show that the expression inside the big brackets is non-negative, it is convenient to define the following four scalars

$$a := v_1^\top p_0, \quad b := v_1^\top p_2, \quad c := v_2^\top p_0, \quad d := v_2^\top p_1,$$

and rewrite that expression as:

$$(a-b)^2 + (c-d)^2 + (a-d)^2 + (b-c)^2 - 2ac + 4bd. \tag{106}$$

First assume that $a = b = 0$. In this situation, (106) reduces to

$$(c-d)^2 + d^2 + c^2, \tag{107}$$

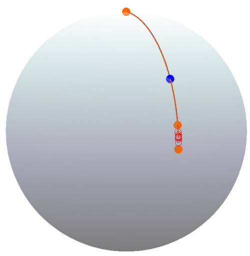
which is evidently nonnegative. Now, assume that $a \neq 0$ or $b \neq 0$. After doing some extensive yet straightforward computations, (106) can be rewritten as

$$\tfrac{3(bc+a(d-c))^2 + (ad+bc-2bd+ca-2(a^2+b^2-ab))^2}{2(a^2+b^2-ab)},$$

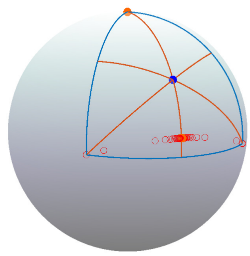
which is nonnegative, since $a^2 + b^2 - ab = ((a-b)^2 + a^2 + b^2)/2$.

In conclusion, the Riemannian Hessian is positive semidefinite and therefore the critical point in the second case is indeed a point of local minimum. □

Figure 6 shows the convergence behavior of points (p_1, p_2) in a neighborhood of a saddle point. If p_1 and p_2 belong to the geodesic that contains p_0, then they converge to the saddle point. Otherwise, they converge to the minimum of the cost function (the vertices of a spherical equilateral triangle).



(a) $p_1 = (0.9854, 0, 0.017)$, $p_2 = (0.9996, 0, -0.0292)$



(b) $p_1 = (0.9975, 0.01, 0.0707)$, $p_2 = (0.9975, -0.01, 0.0707)$

Figure 6: Behavior of a neighborhood of a saddle point.

Acknowledgements

L. Machado and F. Silva Leite are grateful to Fundação para a Ciência e Tecnologia (FCT) for the financial support to the project UIDB/00048/2020 (https://doi.org/10.54499/UIDB/00048/2020). K. Hüper has been supported by the German Federal Ministry of Education and Research (BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA). K. Krakowski and F. Silva Leite thank the Cardinal Stefan Wyszyński University in Warsaw for financial support.

References

[1] Absil, P.A., Mahony, R., Sepulchre, R., 2008. Optimization algorithms on matrix manifolds. Princeton University Press.

[2] Absil, P.A., Mahony, R., Trumpf, J., 2013. An extrinsic look at the riemannian hessian, in: Nielsen, F., Barbaresco, F. (Eds.), Geometric Science of Information, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 361–368.

[3] Afsari, B., 2011. Riemannian L^p center of mass: existence, uniqueness, and convexity. Proc. Amer. Math. Soc. 139, 655–673. URL: `http://dx.doi.org/10.1090/S0002-9939-2010-10541-5`, doi:`10.1090/S0002-9939-2010-10541-5`.

[4] Ando, T., Li, C.K., Mathias, R., 2004. Geometric means. Linear Algebra and its Applications 385, 305–334. URL: https://www.sciencedirect.com/science/article/pii/S0024379503008693, doi:https://doi.org/10.1016/j.laa.2003.11.019. special Issue in honor of Peter Lancaster.

[5] Bačák, M., 2014. Computing medians and means in Hadamard spaces. SIAM J. Optim. 24, 1542–1566. URL: https://doi.org/10.1137/140953393, doi:10.1137/140953393.

[6] Bhatia, R., Holbrook, J., 2006. Riemannian geometry and matrix geometric means. Linear Algebra and its Applications 413, 594–618.

[7] Buss, S.R., Fillmore, J.P., 2001. Spherical averages and applications to spherical splines and interpolation. ACM Trans. Graph. 20, 95–126. doi:http://doi.acm.org/10.1145/502122.502124.

[8] Conway, J.H., Hardin, R.H., Sloane, N.J.A., 1996. Packing lines, planes, etc.: packings in Grassmannian spaces. Experiment. Math. 5, 139–159.

[9] Gallier, J., Quaintance, J., 2020. Differential Geometry and Lie Groups: A Computational Perspective. Springer. doi:10.1007/978-3-030-46040-2.

[10] Hairer, E., Lubich, C., Wanner, G., 2006. Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations. 2nd ed. Springer Series in Computational Mathematics 31. Berlin: Springer.

[11] Horn, R.A., Johnson, C.R., 2013. Matrix analysis. Second ed., Cambridge University Press, Cambridge.

[12] Hüper, K., Silva Leite, F., 2023. Endpoint Geodesic Formulas on Graßmannians Applied to Interpolation Problems. Mathematics 11, (16) 3545; https://doi.org/10.3390/math11163545.

[13] Karcher, H., 1977. Riemannian center of mass and mollifier smoothing. Communications on Pure and Appl. Math. XXX, 509–541.

[14] Karcher, H., 2014. Riemannian center of mass and so called karcher mean URL: https://arxiv.org/abs/1407.2087, doi:10.48550/ARXIV.1407.2087.

[15] Kendall, W.S., 1990. Probability, convexity, and harmonic maps with small image. I: Uniqueness and fine existence. Proc. Lond. Math. Soc., III. Ser. 61, 371–406. doi:10.1112/plms/s3-61.2.371.

[16] Krakowski, K., 2002. Geometrical Methods for Inference. Ph.D. thesis. Department of Mathematics and Statistics. The University of Western Australia.

[17] Lang, S., 1998. Fundamentals of Riemannian Geometry. Springer-Verlag, New York, NY.

[18] Lawson, J., Lim, Y., 2014. Karcher means and Karcher equations of positive definite operators. Trans. Amer. Math. Soc. Ser. B 1, 1–22. URL: https://doi.org/10.1090/S2330-0000-2014-00003-4, doi:10.1090/S2330-0000-2014-00003-4.

[19] Lee, J.M., 2018. Introduction to Riemannian manifolds. 2^{nd} ed., Springer, Cham, Switzerland.

[20] Machado, L., 2006. Least squares problems on Riemannian manifolds. Ph.D. thesis. Department of Mathematics. University of Coimbra, Portugal.

[21] Meyer, C.D., 2000. Matrix analysis and applied linear algebra (incl. CD-ROM and solutions manual). Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics.

[22] Miller, K.S., 1981. On the inverse of the sum of matrices. Mathematics Magazine 54, 67–72.

[23] Moakher, M., 2005. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM Journal on Matrix Analysis and Applications 26, 735–747.

[24] Pennec, X., 2006. Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. Journal of Mathematical Imaging and Vision 25, 127–154. URL: http://dx.doi.org/10.1007/s10851-006-6228-4, doi:10.1007/s10851-006-6228-4.