

# Barycentric Subspaces and Affine Spans in Manifolds

Xavier Pennec<sup>1</sup>

Inria Sophia-Antipolis and Côte d’Azur University (UCA), France  
xavier.pennec@inria.fr

**Abstract.** This paper addresses the generalization of Principal Component Analysis (PCA) to Riemannian manifolds. Current methods like Principal Geodesic Analysis (PGA) and Geodesic PCA (GPCA) minimize the distance to a "Geodesic subspace". This allows to build sequences of nested subspaces which are consistent with a forward component analysis approach. However, these methods cannot be adapted to a backward analysis and they are not symmetric in the parametrization of the subspaces. We propose in this paper a new and more general type of family of subspaces in manifolds: *barycentric subspaces* are implicitly defined as the locus of points which are weighted means of  $k+1$  reference points. Depending on the generalization of the mean that we use, we obtain the Fréchet / Karcher barycentric subspaces (FBS / KBS) or the affine span (with exponential barycenter). This definition restores the full symmetry between all parameters of the subspaces, contrarily to the geodesic subspaces which intrinsically privilege one point. We show that this definition defines locally a submanifold of dimension  $k$  and that it generalizes in some sense geodesic subspaces. Like PGA, barycentric subspaces allow the construction of a forward nested sequence of subspaces which contains the Fréchet mean. However, the definition also allows the construction of backward nested sequence which may not contain the mean. As this definition relies on points and do not explicitly refer to tangent vectors, it can be extended to non Riemannian geodesic spaces. For instance, principal subspaces may naturally span over several strata in stratified spaces, which is not the case with more classical generalizations of PCA.

## 1 Introduction

For Principal Component Analysis (PCA) in a Euclidean space, one can equivalently define the principal  $k$ -dimensional affine subspace using the minimization of the variance of the residuals (the projection of the data point to the subspace) or the maximization of the explained variance within that affine subspace. This is due to the Pythagorean theorem, which does not hold in more general manifolds. A second important observation is that principal components of different orders are nested, which allows to build forward and backward estimation methods by iteratively adding or removing principle components.

Generalizing affine subspaces to manifolds is not so obvious. For the zero-dimensional subspace, intrinsic generalization of the mean on manifolds naturally comes into mind: the Fréchet mean is the set of global minima of the variance, as defined by Fréchet in general metric spaces [5]. The set of local minima of the variance was named Karcher mean by W.S Kendall [10] after the work of Karcher et al. on Riemannian centers of mass ([8] see [9] for a discussion of the naming and earlier works).

The one-dimensional component is then quite naturally a geodesic which should pass through the mean point. Higher-order components are more difficult to define. The simplest intrinsic generalization of PCA is tangent PCA (tPCA), which amounts to unfold the whole distribution in the tangent space at the mean using the pullback of the Riemannian exponential map, and to compute the principle components of the covariance matrix in the tangent space. The method is thus based on the maximization of the explained variance. tPCA is often used on manifolds because it is simple and efficient. However, if it is good for analyzing data which are sufficiently centered around a central value (unimodal or Gaussian-like data), it is often not sufficient for multimodal or large support distributions (e.g. uniform on close compact subspaces).

Fletcher et al. proposed in [4] to rely on the least square distance to subspaces which are totally geodesic at one point. These Geodesic Subspaces (GS) are spanned by the geodesics going through one point with tangent vector restricted to a linear subspace of the tangent space. These subspaces are only locally a manifold as they are generally not smooth at the cut locus of the mean point. The procedure was coined Principle Geodesic Analysis (PGA). However, the least-square procedure was computationally expensive, so that the authors implemented in practice a classical tangent PCA. A real implementation of the original PGA procedure was only provided recently by Sommer et al. [16]. PGA is intrinsic and allows to build a sequences of embedded principal geodesic subspaces in a forward component analysis approach by building iteratively the components from dimension 0 (the mean point), dimension 1 (a geodesic), etc. Higher dimensions are obtained iteratively by selecting the direction in the tangent space at the mean that optimally reduce the square distance of data point to the geodesic subspace. However, the mean always belong to geodesic subspaces even when it is not part of the support of the distribution.

Huckemann et al. [14] proposed to start at the first order component by fitting a geodesic to the data, not necessarily through the mean. The second principle geodesic is chosen orthogonally to the first one, and higher order components are added orthogonally at the crossing point to build a geodesic subspace. The method was named Geodesic PCA (GPCA). Sommer [15] proposed a method called horizontal component analysis (HCA) which uses the parallel transport of the 2nd direction along the first principle geodesic to define the second coordinates, and iteratively define higher order coordinates through horizontal development along the previous modes. Other principle decompositions have been proposed, like Principle Graphs [6], extending the idea of k-means.

All the cited methods are intrinsically forward methods that build successively larger approximation spaces for the data. A notable exception is Principle Nested Spheres (PNS), proposed by Jung, et al. [7] as a general framework for non-geodesic decomposition of high-dimensional spheres or high-dimensional planar landmarks shape spaces. Subsphere or radius 0 to 1 are obtained by slicing a higher dimensional sphere by an affine hyperplane. The backward analysis approach, determining a decreasing family of subspace, has been generalized to more general manifold with the help of a nested sequence of relations [3]. However, up to know, such sequences of relationships are only known for spheres, Euclidean spaces or quotient spaces of Lie groups by isometric actions [14].

In this paper, we keep the principle of minimizing the unexplained information. However, we propose to replace Geodesic Subspaces by new and more general types of family of subspaces in manifolds: Barycentric Subspaces (BS). BS are defined as the locus of points which are weighted means of  $k + 1$  reference points. Depending on the generalization of the mean that we use on manifolds, Fréchet mean, Karcher mean or exponential barycenter, we obtain the Fréchet / Karcher barycentric subspaces (FBS / KBS) or the affine span. We show that these definition are related and locally define a submanifold of dimension  $k$ , and that they generalize in some sense the geodesic subspaces. Like PGA, Barycentric Subspace Analysis (BSA) allows the construction of forward nested subspaces which contains the Fréchet mean. However, it also allows a backward analysis which may not contain the mean. As this definition relies on points and do not explicitly refer to tangent vectors to parametrize geodesics, a very interesting side effect is that it can also be extended to more general geodesic spaces that are not Riemannian. For instance, in stratified spaces, it naturally allows to have principle subspaces that span over several strata. The paper is divided in three parts. We recall in Section 2 the background knowledge. Then, we define in Section 3 the notions of barycentric subspaces in metric spaces and the affine spans in manifolds. Section 4 finally establishes important properties and relationships between these subspaces.

## 2 Background knowledge on Riemannian manifolds

### 2.1 Computing in Riemannian manifolds

We consider an embedding Riemannian manifold  $\mathcal{M}$  of dimension  $n$ . The Riemannian metric is denoted  $\langle \cdot | \cdot \rangle_x$  on each tangent space  $T_x\mathcal{M}$  of the manifold. The expression of the the underlying norm in a chart is  $\|v\|_x^2 = v^T G(x) v = v^i v^j g_{ij}(x)$  using Einstein notations for tensor contractions. We assume the manifold to be geodesically complete (no boundary nor any singular point that we can reach in a finite time). As an important consequence, the Hopf-Rinow-De Rham theorem states that there always exists at least one minimizing geodesic between any two points of the manifold.

We denote by  $\exp_x(v)$  the *exponential map* at point  $x$  which associate to each tangent vector  $v \in T_x\mathcal{M}$  the point of  $\mathcal{M}$  reached by the geodesic starting at  $x$  with this tangent vector after a unit time. This map is a local diffeomorphism

from  $0 \in T_x\mathcal{M}$  to  $\mathcal{M}$ , and we denote  $\vec{x}\vec{y} = \log_x(y)$  its inverse: it may be defined as the smallest vector of  $T_x\mathcal{M}$  that allows to shoot a geodesic from  $x$  to  $y$ . A geodesic  $\exp_x(tv)$  is minimizing up to a certain cut time  $t_0$  and not anymore after. When  $t_0$  is finite,  $t_0v$  is called a tangential cut-point and  $\exp_x(t_0v)$  a cut point. The domain of injectivity  $D(x) \in T_x\mathcal{M}$  of the exponential map can be maximally extended up to the tangential cut-locus  $\partial D(x) = C(x)$ . It covers all the manifold  $\mathcal{M}$  except the *cut locus*  $\mathcal{C}(x) = \exp_x(C(x))$  which has null measure for the Riemannian measure.

When the tangent space is provided with an orthonormal basis, the Riemannian exponential and logarithmic maps provide a *normal coordinate systems at  $x$* . A set of normal coordinate systems at each point of the manifold realize an atlas which allows to work very easily on the manifold. The implementation of exp and log maps is the basis of programming on Riemannian manifolds, and most the geometric operations needed for statistics or image processing can be rephrased based on them [12, 13].

## 2.2 Taylor expansions in normal coordinate systems

We consider a normal coordinate system centered at  $x$  and  $x_v = \exp_x(v)$  a variation of the point  $x$ . We denote by  $R_{jkl}^i(x)$  the coefficients of the Riemannian curvature tensor at  $x$  and by  $\epsilon$  a conformal gauge scale that encodes the size of the path (in terms of  $\|v\|_x$  and  $\|\vec{x}\vec{y}\|_x$ ) normalized by the curvature. Following [2], the Taylor expansion of the metric is  $g_b^a(v) = \delta_b^a - \frac{1}{3}R_{cbd}^a v^c v^d - \frac{1}{6}\nabla_e R_{cbd}^a v^e v^c v^d + O(\epsilon^4)$ , and a geodesic joining  $x_v$  to  $y$  has initial tangent vector:

$$[\log_{x_v}(y)]^a = \vec{x}\vec{y}^a - v^a + \frac{1}{3}R_{cbd}^a v^b \vec{x}\vec{y}^c \vec{x}\vec{y}^d + \frac{1}{12}\nabla_c R_{dbe}^a v^b \vec{x}\vec{y}^c \vec{x}\vec{y}^d \vec{x}\vec{y}^e + O(\epsilon^4).$$

Combining these two expansions, we get the expansion of the Riemannian distance:  $d_{xy}^2(v) = \text{dist}^2(\exp_x(v), y) = \|\vec{x}\vec{y}\|_x^2 + (\nabla d_{xy}^2)^T v + \frac{1}{2}v^T \nabla^2 d_{xy}^2 v + O(\epsilon^3)$ , where the gradient  $\nabla d_{xy}^2 = -2\vec{x}\vec{y}$  is -2 times the log and the Hessian is the opposite of the differential of the log:

$$(\nabla^2 d_{xy}^2)_b^a = -[D_x \log_x(y)]_b^a = \delta_b^a - \frac{1}{3}\vec{x}\vec{y}^c \vec{x}\vec{y}^d R_{cbd}^a - \frac{1}{12}\vec{x}\vec{y}^c \vec{x}\vec{y}^d \vec{x}\vec{y}^e \nabla_c R_{dbe}^a + O(\epsilon^3).$$

## 2.3 Moments of point distributions

Let  $\mu(x) = \sum_i \lambda_i \delta_{x_i}(x)$  be a singular distribution of  $k+1$  points on  $\mathcal{M}$  with weights  $(\lambda_0, \dots, \lambda_k)$  that do not sum up to zero. To define the moments of that distribution, we have to take care that the Riemannian log and distance functions are not smooth at the cut-locus of the points  $\{x_i\}$ .

### Definition 1 (( $k+1$ )-pointed Riemannian manifold).

Let  $\{x_0, \dots, x_k\} \in \mathcal{M}^{k+1}$  be a set of  $k+1$  distinct points in the Riemannian manifold  $\mathcal{M}$  and  $C(x_0, \dots, x_k) = \cup_{i=0}^k C(x_i)$  be the union of the cut loci of these points. We call ( $k+1$ )-pointed manifold  $\mathcal{M}^*(x_0, \dots, x_k) = \mathcal{M}/C(x_0, \dots, x_k)$  the submanifold of the non-cut points of the points.

Since the cut locus of each point is closed and has null measure,  $\mathcal{M}^*(x_0, \dots, x_k)$  is open and dense in  $\mathcal{M}$ . Thus, it is a submanifold of  $\mathcal{M}$  (not necessarily connected). On this submanifold  $\mathcal{M}^*(x_0, \dots, x_k)$ , the distance to the points  $x_i$  and the Riemannian log function  $\vec{x}\vec{x}_i = \log_x(x_i)$  are smooth.

**Definition 2 (Weighted moments of a  $(k + 1)$ -pointed manifold).**

Let  $(\lambda_0, \dots, \lambda_k) \in \mathbb{R}^{k+1}$  such that  $\sum_i \lambda_i \neq 0$ . The weighted  $n$ -order moment of a  $(k + 1)$ -pointed Riemannian manifold  $\mathcal{M}^*(x_0, \dots, x_k)$  is the smooth  $(n, 0)$  tensor:

$$\mathfrak{M}_n(x, \lambda) = \sum_i \lambda_i \underbrace{\overrightarrow{xx_i} \otimes \overrightarrow{xx_i} \dots \otimes \overrightarrow{xx_i}}_{n \text{ times}} \quad (1)$$

The 0-th order moment (the mass)  $\mathfrak{M}_0(\lambda) = \sum_i \lambda_i = \mathbf{1}^\top \lambda$  is constant. All other moments are homogeneous of degree 1 in  $\lambda$  and can be normalized by dividing by the mass  $\mathfrak{M}_0(\lambda)$ . The first order moment  $\mathfrak{M}_1(x, \lambda) = \sum_i \lambda_i \overrightarrow{xx_i}$  is a smooth vector field on the manifold  $\mathcal{M}^*(x_0, \dots, x_k)$ . The second and higher order moments are smooth  $(n, 0)$  tensor fields that will be used through their contraction with the Riemannian curvature tensor.

### 3 Barycentric subspaces

In a Euclidean space, an affine subspace of dimension  $k$  is generated by a point and  $k$  non-collinear vectors:  $\text{Aff}(x_0, v_1 \dots v_k) = \left\{ x = x_0 + \sum_{i=1}^k \lambda_i v_i, \lambda \in \mathbb{R}^k \right\}$ . Alternatively, one could also generate the affine span of  $k + 1$  points in general linear position using the implicit equation  $\sum_i \lambda_i (x_i - x) = 0$  where  $\sum_{i=0}^k \lambda_i = 1$ . The two definitions are equivalent when  $x_i = x_0 + v_i$ . The last parametrization of  $k$ -dimensional affine submanifolds is relying on barycentric coordinates which live in the projective space  $\mathcal{P}_k$  minus the orthogonal of the line element  $\mathbf{1} = (1 : 1 : \dots : 1)$ :

$$\mathcal{P}_k^* = \left\{ (\lambda_0 : \dots : \lambda_k) \in \mathbb{R}^{k+1} \text{ s.t. } \sum_i \lambda_i \neq 0 \right\}.$$

Standard charts of this space are given either by the intersection of the line elements with the "upper" unit sphere  $S_k$  of  $\mathbb{R}^{k+1}$  with north pole  $\mathbf{1}/\sqrt{k}$  (unit weights) or by the  $k$ -plane of  $\mathbb{R}^{k+1}$  passing through the point  $\mathbf{1}/k$  and orthogonal to this vector. We call normalized weights  $\underline{\lambda}_i = \lambda_i / (\sum_{j=0}^k \lambda_j)$  this last projection.

#### 3.1 Fréchet and Karcher Barycentric subspaces in a metric space

The two above definitions of the affine span turn out to have different generalizations in manifolds: the first definition leads to geodesic subspaces, as defined in PGA and GPCA [4, 16, 14], while the second definition using the affine span suggests a generalization to manifolds either using the Fréchet/Karcher weighted mean or using an exponential barycenter.

**Definition 3 (Fréchet / Karcher Barycentric subspaces of  $k+1$  points).**

Let  $(\mathcal{M}, \text{dist})$  be a metric space and  $(x_0, \dots, x_k) \in \mathcal{M}^k$  be  $k + 1$  distinct reference points. The (normalized) weighted variance at point  $x$  with weight  $\lambda \in \mathcal{P}_k^*$  is:  $\sigma^2(x, \lambda) = \frac{1}{2} \sum_{i=0}^k \underline{\lambda}_i \text{dist}^2(x, x_i) = \frac{1}{2} \sum_{i=0}^k \lambda_i \text{dist}^2(x, x_i) / (\sum_{j=0}^k \lambda_j)$ . The Fréchet barycentric subspace is the locus of weighted Fréchet means of these

points, i.e. the set of absolute minima of the weighted variance:  $FBS(x_0, \dots, x_k) = \{\arg \min_{x \in \mathcal{M}} \sigma^2(x, \lambda), \lambda \in \mathcal{P}_k^*\}$ . The Karcher barycentric subspace  $KBS(x_0, \dots, x_k)$  is defined similarly with local minima instead of global ones.

This definition restores the full symmetry of all the parameters defining the subspaces, contrarily to the geodesic subspaces which privilege one point. Here, we defined the notion on general metric spaces to show that it works in spaces more general than smooth Riemannian manifolds. In a stratified space for instance, the barycentric subspace spanned by points belonging to different strata naturally maps over all these strata. This is a significant improvement over geodesic subspaces used in PGA which can only be defined within a regular strata.

### 3.2 Affine spans as exponential barycentric subspaces

A second way to generalize the affine span to manifolds is to see directly the implicit barycentric coordinates equation as a weighted exponential barycenter:

**Definition 4 (Affine span of a  $(k + 1)$ -pointed Riemannian manifold).**

A point  $x \in \mathcal{M}^*(x_0, \dots, x_k)$  has barycentric coordinates  $\lambda \in \mathcal{P}_k^*$  if

$$\mathfrak{M}_1(x, \lambda) = \sum_{i=0}^k \lambda_i \overrightarrow{xx_i} = 0. \quad (2)$$

The affine span of the points  $(x_0, \dots, x_k) \in \mathcal{M}^k$  is the set of weighted exponential barycenters of the reference points in  $\mathcal{M}^*(x_0, \dots, x_k)$ :

$$\text{Aff}(x_0, \dots, x_k) = \{x \in \mathcal{M}^*(x_0, \dots, x_k) \mid \exists \lambda \in \mathcal{P}_k^* : \mathfrak{M}_1(x, \lambda) = 0\}.$$

This definition is only valid on  $\mathcal{M}^*(x_0, \dots, x_k)$  and may hide some discontinuities of the affine span on the union of the cut locus of the reference points. Outside this null measure set, one recognizes that Eq.(2) defines nothing else than the critical points of the variance  $\sigma^2(x, \lambda) = \frac{1}{2} \sum_i \lambda_i \text{dist}^2(x, x_i)$ . The affine span is thus a superset of the barycentric subspaces in  $\mathcal{M}^*(x_0, \dots, x_k)$ . However, we notice that the variance may also have local minima on the cut-locus of the reference points.

Let us consider field of  $n \times (k + 1)$  matrices  $Z(x) = [\overrightarrow{xx_0}, \dots, \overrightarrow{xx_k}]$ . We can rewrite Eq.(2) in matrix form:  $\mathfrak{M}_1(x, \lambda) = Z(x)\lambda = 0$ . Thus, we see that the affine span is controlled by the kernel of the matrix field  $Z(x)$ :

**Theorem 1 (SVD Characterization of the affine span).**

Let  $Z(x) = U(x).S(x).V(x)^T$  be a singular decomposition of the matrix fields  $Z(x) = [\overrightarrow{xx_0}, \dots, \overrightarrow{xx_k}]$  on  $\mathcal{M}^*(x_0, \dots, x_k)$  (with singular values sorted in decreasing order). The barycentric subspace  $\text{Aff}(x_0, \dots, x_k)$  is the zero level-set of the  $k + 1$  singular value  $s_{k+1}(x)$  and the subspace of valid barycentric weights is spanned by the right singular vectors corresponding to the  $l$  vanishing singular values:  $\text{Span}(v_{k-l}, \dots, v_k)$  (it is void if  $l = 0$ ).

## 4 Properties of barycentric subspaces in manifolds

In this section, we restrict the analysis to  $\mathcal{M}^*(x_0, \dots, x_k)$  so that all quantities are smooth.

#### 4.1 Karcher barycentric subspaces and affine span

In  $\mathcal{M}^*(x_0, \dots, x_k)$ , the critical points of the weighted variance are the points of the affine span. Among these points, the local minima may be characterized by the Hessian  $H(x, \lambda) = -\sum_i \lambda_i D_x \log_x(x_i)$  of the weighted variance. Using the Taylor expansion of the differential of the log of Section 2.2), we obtain:

$$[H(x, \lambda)]_b^a = \delta_b^a - \frac{1}{3} R_{cbd}^a [\mathfrak{M}_2(x, \lambda)]^{cd} - \frac{1}{12} \nabla_c R_{dbe}^a [\mathfrak{M}_3(x, \lambda)]^{cde} + O(\epsilon^4), \quad (3)$$

The key factor is the contraction of the curvature with the dispersion of the reference points: when the typically distance from  $x$  to all the reference points  $x_i$  is smaller than the inverse of the curvature, then  $H(x, \lambda)$  is essentially close to the identity. In the limit of null curvature, (e.g. for a Euclidean space),  $H(x, \lambda)$  is simply the unit matrix. In general Riemannian manifolds, Eq.(3) only gives a qualitative behavior. In order to obtain hard bounds on the spectrum of  $H(x, \lambda)$ , one has to investigate bounds on Jacobi fields, as is done for the proof of uniqueness of the Karcher and Fréchet means [8, 10, 11, 1, 17]. Thanks to these proofs, we can in fact establish that the Karcher barycentric submanifold is locally well defined around the Karcher mean.

When the Hessian is degenerated, we cannot conclude on the local minimality without going to higher order differentials. This leads us to stratify the affine span by the index of the Hessian of the weighted variance.

**Definition 5 (Regular and positive points of  $\mathcal{M}^*(x_0, \dots, x_k)$ ).**

*A point  $x \in \mathcal{M}^*(x_0, \dots, x_k)$  is said regular (resp. positive) if the Hessian matrix  $H(x, \lambda)$  is invertible (resp. positive definite) for all  $\lambda$  in the right singular space of the smallest singular value of  $Z(x)$ . The set of regular (resp. positive) points is denoted  $\text{Reg}(\mathcal{M}^*(x_0, \dots, x_k))$  (resp.  $\text{Reg}^+(\mathcal{M}^*(x_0, \dots, x_k))$ ). The set of positive points of the affine span is called the positive span  $\text{Aff}^+(x_0, \dots, x_k)$ .*

Positive points are obviously regular, and in Euclidean spaces all the points are positive and regular. However, in Riemannian manifolds, we may have non-regular points and regular points which are non-positive.

**Theorem 2 (Karcher barycentric subspace and positive span).**

*The positive span  $\text{Aff}^+(x_0, \dots, x_k)$  is the set of regular points of the Karcher barycentric subspace  $KBS(x_0, \dots, x_k)$  on  $\mathcal{M}^*(x_0, \dots, x_k)$ .*

One generalization of the Fréchet (resp. Karcher) mean is the use of the power  $\alpha$  of the metric instead of the square. For instance, one defines the median ( $\alpha = 1$ ) and the modes ( $\alpha \rightarrow 0$ ) as the minima of the  $\alpha$ -variance  $\sigma^\alpha(x) = \frac{1}{\alpha} \sum_{i=0}^k \text{dist}^\alpha(x, x_i)$ . Following this idea, one could think of generalizing barycentric subspaces to the  $\alpha$ -Fréchet (resp.  $\alpha$ -Karcher) barycentric subspaces. In fact, it turns out that the critical points of the  $\alpha$ -variance are just elements of the affine span with weights  $\lambda'_i = \lambda_i \text{dist}^{\alpha-2}(x, x_i)$ . Thus, changing the power of the metric just amounts to reparametrizing the barycentric weights, which shows the notion of affine span is really central.

## 4.2 Dimension of the barycentric subspace

We can locally parametrize the affine span thanks to a Taylor expansion of the constraint  $Z(x)\lambda = 0$ : a change of coordinates  $\delta\lambda$  induces a change of position  $\delta x$  verifying  $H(x, \lambda)\delta x + Z(x)\delta\lambda = 0$ . At the positive points, the Hessian is invertible and the SVD characterization leads us to conclude that:

**Theorem 3 (Dimension of the barycentric subspaces at regular points).** *The positive span  $\text{Aff}^+(x_0, \dots, x_k)$  (i.e. the regular KBS), is a stratified space of dimension  $k$  on  $\text{Reg}(\mathcal{M}^*(x_0, \dots, x_k))$ . On the  $m$ -dimensional strata,  $Z(x)$  has exactly  $k - m + 1$  vanishing singular values.*

## 4.3 Geodesic subspaces as limit of barycentric subspaces

By analogy with Euclidean spaces, one would expect the affine span to be close to the geodesic subspace

$$GS(x, w_1, \dots, w_k) = \left\{ \exp_x \left( \sum_{i=1}^k \alpha_i w_i \right) \in \mathcal{M} \text{ for } \alpha \in \mathbb{R}^k \right\}$$

generated by the  $k$  independent vectors  $w_1, \dots, w_k$  at  $x$  when all the points  $\{x_i = \exp_{x_0}(\epsilon w_i)\}_{1 \leq i \leq k}$  are converging to  $x_0$  at first order.

In order to investigate that, we first need to restrict the definition of the geodesic subspaces. Indeed, although the above classical definition is implicitly used in most of the works using PGA, it may not define a  $k$ -dimensional submanifold when there is a cut-locus. For instance, it is well known that geodesics of a flat square torus are either periodic or everywhere dense in a flat torus submanifold depending on whether the components of the initial velocity field have rational or irrational ratios. Thus, it makes sense to restrict to the part of the GS which is limited by the cut-locus.

### Definition 6 (Restricted Geodesic Submanifolds).

Let  $x \in \mathcal{M}$  be a point of a Riemannian manifold and  $W_x = \left\{ \sum_{i=1}^k \alpha_i w_i, \alpha \in \mathbb{R}^k \right\}$  the  $k$ -dimensional linear subspace of  $T_x \mathcal{M}$  generated a  $k$ -uplet  $\{w_i\}_{1 \leq i \leq k} \in (T_x \mathcal{M})^k$  of tangent vectors at  $x$ . Recall that  $D(x) \subset T_x \mathcal{M}$  is the maximal definition domain on which the exponential map is diffeomorphic.

We call restricted geodesic submanifold  $GS^*(W_x)$  at  $x$  generated by the vector subspace  $W_x$  the submanifold of  $\mathcal{M}$  generated by the geodesics starting at  $x$  with tangent vectors  $w \in W_x$ , but up to the first cut-point of  $x$  only:

$$GS^*(W_x) = GS^*(x, w_1, \dots, w_k) = \{ \exp_x(w), w \in W_x \cap D(x) \}$$

This restricted definition correctly defines a  $k$ -dimensional submanifold of  $\mathcal{M}$ , whose completion may be a manifold with boundary.

Let  $x = \exp_{x_0}(w) \in GS^*(W_x)$ . Thanks to the symmetry of geodesics, we can show that this point is solution of the barycentric equation  $\sum_{i=0}^k \lambda_i \log_x(x_i) = O(\eta^2)$  with non-normalized homogeneous coordinates  $\lambda_i = \alpha_i$  for  $1 \leq i \leq k$  and  $\lambda_0 = \eta - (\sum_i \alpha_i)$ . These coordinates obviously sum up to zero when  $\eta$



goes to zero, which is a point at infinity in  $\mathcal{P}_k^*$ . In that sense, points of the restricted geodesic submanifold  $GS^*(W)$  are points at infinity of the affine span  $\text{Aff}(x, x_1, \dots, x_k)$  when the points  $x_i = \exp_x(\eta w_i)$  are converging to  $x$  at first order along the tangent vectors  $w_i$ .

**Theorem 4 (Restricted GS as limit case of the affine span).**

*Points of the restricted geodesic submanifold  $GS^*(W_x) = \{\exp_x(w), w \in W_x \cap D(x)\}$  are points at infinity in  $\mathcal{P}_k^*$  of the affine span  $\text{Aff}(x, x_1, \dots, x_k)$  when the points  $x_i = \exp_x(\eta w_i)$  are converging to  $x$  at first order along the tangent vectors  $w_i$  defining the  $k$ -dimensional subspace  $W_x \subset T_x\mathcal{M}$ .*

## 5 Perspectives

We proposed in this paper three generalization of the affine span of  $k + 1$  points in a manifold. These barycentric subspaces are implicitly defined as the locus of points which are weighted (Fréchet / Karcher / exponential barycenter) means of  $k + 1$  reference points. In generic conditions, barycentric subspaces are stratified spaces that are locally submanifolds of dimension  $k$ . Their singular set of dimension  $k - l$  corresponds to the case where  $l$  of the reference point belongs to the barycentric subspace defined by the  $k - l$  other reference points.

In non-generic conditions, points may coalesce along certain directions, defining non local jets instead of a regular  $k$ -tuple. Geodesic subspaces, which are defined by  $k - 1$  tangent vectors at a point, do correspond (in some restricted sense) to the limit of the affine span when the  $k$ -tuple converges towards that jet. We conjecture that this can be generalized to higher order derivatives using techniques from sub-Riemannian geometry. This way, some non-geodesic decomposition schemes such as loxodromes, splines and principle nested spheres could also be seen as limit cases of barycentric subspaces.

Investigating simple manifolds like spheres and symmetric spaces will provide useful guidelines in that direction. For instance, the closure of the barycentric subspace of  $k + 1$  different reference points on the  $n$ -dimensional sphere is the  $k$ -dimensional great subsphere that contains the reference points. It is noticeable that the closure of the affine span generated by any  $k + 1$ -tuple of points of a great  $k$ -dimensional subsphere generate the same space, which is also a geodesic subspace. This coincidence of spaces is due to the very high symmetry of the sphere. For second order jets, we conjecture that we obtain subspheres of different radii as used in principle nested spheres (PNS) analysis.

Barycentric subspaces can be naturally nested, by defining an ordering of the reference points, which makes it suitable for a generalization of Principal Component Analysis (PCA) to Riemannian manifolds. Several problems however remain to be investigated to use Barycentric Subspace Analysis (BSA) in practice. First, the optimization on  $k$ -tuple might have multiple solutions, as in the case of spheres. Here, we need to find a suitable quotient space similar to the quotient definition of Grassmannians. Second, the optimization might converge towards a non-local jet instead on a  $k$ -tuple, and good renormalization

techniques need to be designed to guaranty the numerical stability. Third, one theoretically needs to define a proper criterion to be optimized by all  $k$ -tuple for  $k = 0 \dots n$  together and not just a greedy approach as done by the classical forward and backward approaches.

## References

1. B. Afsari. Riemannian  $l^p$  center of mass: existence, uniqueness, and convexity. *Proc. of the AMS*, 180(2):655–673, Feb. 2010.
2. L. Brewin. Riemann normal coordinate expansions using cadabra. *Classical and Quantum Gravity*, 26(17):175017, 2009.
3. J. Damon and J. S. Marron. Backwards Principal Component Analysis and Principal Nested Relations. *Journal of Mathematical Imaging and Vision*, 50(1-2):107–114, Oct. 2013.
4. P. Fletcher, C. Lu, S. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005, Aug. 2004.
5. M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'Institut Henri Poincaré*, 10:215–310, 1948.
6. A. N. Gorban and A. Y. Zinovyev. Principal graphs and manifolds. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques*, chapter 2, pages 28–59. 2009.
7. S. Jung, I. L. Dryden, and J. S. Marron. Analysis of principal nested spheres. *Biometrika*, 99(3):551–568, Sept. 2012.
8. H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications in Pure and Applied Mathematics*, 30:509–541, 1977.
9. H. Karcher. Riemannian Center of Mass and so called karcher mean. *arXiv:1407.2087 [math]*, July 2014. arXiv: 1407.2087.
10. W. Kendall. Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence. *Proc. London Math. Soc.*, 61(2):371–406, 1990.
11. H. Le. Estimation of Riemannian barycenters. *LMS J. Comput. Math.*, 7:193–200, 2004.
12. X. Pennec. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006. A preliminary appeared as INRIA RR-5093, January 2004.
13. X. Pennec, P. Fillard, and N. Ayache. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006. A preliminary version appeared as INRIA Research Report 5255, July 2004.
14. A. M. S. Huckemann, T. Hotz. Intrinsic shape analysis: Geodesic principal component analysis for Riemannian manifolds modulo Lie group actions. *Statistica Sinica*, pages 1–100, 2010.
15. S. Sommer. Horizontal Dimensionality Reduction and Iterated Frame Bundle Development. In F. Nielsen and F. Barbaresco, editors, *Geometric Science of Information*, number 8085 in Lecture Notes in Computer Science, pages 76–83. Springer Berlin Heidelberg, 2013.
16. S. Sommer, F. Lauze, and M. Nielsen. Optimization over geodesics for exact principal geodesic analysis. *Advances in Computational Mathematics*, 40(2):283–313, June 2013.
17. L. Yang. *Medians of probability measures in Riemannian manifolds and applications to radar target detection*. PhD thesis, Poitier University, Dec. 2011.