

Statistica Sinica Preprint No: SS-07-173

Title:	Intrinsic shape analysis: Geodesic PCA for Riemannian manifolds modulo isometric lie group actions
Manuscript ID:	SS-07-173 R2
URL:	http://www.stat.sinica.edu.tw/statistica/
Complete List of Authors:	Stephan Huckemann, Thomas Hotz and Axel Munk
Corresponding Author:	Stephan Huckemann
E-mail:	huckeman@uni-math.gwdg.de

**INTRINSIC SHAPE ANALYSIS:
GEODESIC PCA FOR RIEMANNIAN MANIFOLDS
MODULO ISOMETRIC LIE GROUP ACTIONS**

Stephan Huckemann[†], Thomas Hotz* and Axel Munk

Institute for Mathematical Stochastics, Georgia Augusta Universität Göttingen

Abstract: A general framework is laid out for principal component analysis (PCA) on quotient spaces that result from an isometric Lie group action on a complete Riemannian manifold. If the quotient is a manifold, geodesics on the quotient can be lifted to horizontal geodesics on the original manifold. Thus, PCA on a manifold quotient can be pulled back to the original manifold. In general, however, the quotient space may no longer carry a manifold structure. Still, horizontal geodesics can be well-defined in the general case. This allows for the concept of generalized geodesics and orthogonal projection on the quotient space as the key ingredients for PCA. Generalizing a result of Bhattacharya and Patrangenaru (2003), geodesic scores can be defined outside a null set. Building on that, an algorithmic method to perform PCA on quotient spaces based on generalized geodesics is developed. As a typical example where non-manifold quotients appear, this framework is applied to Kendall's shape spaces. In fact, this work has been motivated by an application occurring in forest biometry where the current method of Euclidean linear approximation is unsuitable for performing PCA. This is illustrated by a data example of individual tree stems whose Kendall shapes fall into regions of high curvature of shape space: PCs obtained by Euclidean approximation fail to reflect between-data distances and thus cannot correctly explain data variation. Similarly, for a classical archeological data set with a large spread in shape space, geodesic PCA allows new insights that have not been available under PCA by Euclidean approximation. We conclude by reporting challenges, outlooks, and possible perspectives of intrinsic shape analysis.

Key words and phrases: Shape Analysis, Principal Component Analysis, Riemannian Manifolds, Orbifolds, Orbit Spaces, Geodesics, Lie Group Actions, Non-Linear Multivariate Statistics, Intrinsic Mean, Extrinsic Mean, Forest Biometry

[†] Supported by DFG Grant MU 1230/10-1 and the Volkswagen Stiftung.

* Supported by DFG Graduiertenkolleg 1023 and by the German Federal Ministry of Education and Research, Grant 03MUPAH6

AMS 2000 Subject Classification: Primary 60D05
Secondary 62H11, 53C22

1 Introduction

In this paper, we illustrate a new approach for applying classical statistical methods to multivariate non-linear data. In two examples occurring in the statistical study of shape of three dimensional geometrical objects, we illustrate that the current methods of PCA by linear Euclidean approximation are unsuitable if such data in non-linear spaces fall into regions of high curvature, or if they have a large spread. In the following we give an overview of the background of relevant previous work, and an introduction to the building blocks of our work.

Euclidean and Non-Euclidean Data. Over the last century, multivariate statistics for *Euclidean* data structures has been the target of intensive research, leading mainly to *linear* statistical methods of analysis. More recently, a growing demand can be observed for methods treating multivariate data on spaces with a natural *non-Euclidean* structure. We mention statistical estimation problems of and on manifolds as they arise in various applications, e.g. Kim and Koo (2005), estimation of manifolds, e.g. de Silva and Carlsson (2004) and Bubenik and Kim (2007), or statistical inference in shape analysis, e.g. Munk et al. (2007), which often require a generalization of the underlying space to a quotient of a manifold, e.g. Kendall et al. (1999), or even more general structures such as semimetrical spaces, e.g. Schmidt et al. (2007). Mostly, such data have been dealt with by linear approximations and quite some advances have been achieved. The aim of this paper is to explore the limitations of such linearizations, and to provide a methodology that may be applied when those linear approximations fail to capture the non-Euclidean nature of the data. We emphasize that we do not claim to solve these issues in full generality, rather that we would like to direct the interest of the readers to this ambitious research program while we restrict our presentation to quotients of manifolds under a Lie group action.

Extrinsic, Euclidean and Intrinsic Methods. A very powerful tool of traditional Euclidian multivariate statistical analysis is *principal component analysis* (PCA). It aims to reduce the dimensionality of the data, and yields a hierarchy of major directions explaining the main sources of data variation. This raises the question of designing a similar tool for data on non-Euclidean spaces. In Table 1 we give an overview of various methods developed in the past and proposed in this paper to tackle that question. Following the idea of linearization this can be done by performing PCA in a Euclidean tangent space (whenever it exists) of an underlying space. Usually, the tangent space at an *extrinsic mean* (EM) is chosen, the latter in the manifold case being an orthogonal projection of the Euclidean mean onto the manifold in an ambient space, e.g. Hendriks, Landsman, and Ruymgaart (1996), as well as Hendriks and Landsman (1998) or, in a more general case, being a *Procrustes mean*, cf. Gower (1975). Often it seems more natural to define an *intrinsic mean* (IM), i.e. a minimizer of the squared intrinsic distance to the data (Kobayashi and Nomizu ((1969), p.109), and Karcher (1977)), where the intrinsic distance is usually determined by the Riemannian structure induced either by the subject matter

or by the specific construction as it is the case for shape spaces (e.g. Le (2001), Bhattacharya and Patrangenaru (2003, 2005), as well as Klassen et al. (2004)). The relationship between EM and IM is not obvious, and not well understood. We mention that in our applications the EM is a fairly good approximation to the IM. Currently, PCA in the tangent space of either mean is performed in a *Euclidean* manner, either by some projection of the data to the tangent space at the EM, or by mapping the data under the inverse Riemann exponential map to the tangent space at the IM. The mapped data serve as the basis for computing the empirical covariance matrix, and hence for PCA. The well known *general Procrustes analysis* (GPA) for quotients, such as Kendall's shape spaces, is based on this procedure by orthogonally projecting the data to the tangent space at an EM, see Gower (1975), Goodall (1991), Cootes et al. (1992) and Kent (1994). Alternatively in *principal geodesic analysis* (PGA), the data is mapped under the inverse Riemann exponential at the IM, see Fletcher et al. (2004). In fact, intrinsic distances between data and mean are equal (under the inverse exponential) or approximately equal (in case of orthogonal projection) to the respective distances in the tangent space. When curvature is present this is not the case for between-data distances, which carry the additional information extracted by PCA. Therefore extrinsic and Euclidean methods, as developed so far, are well suited for statistical analysis focussing on the mean, but may fail to capture the additional information required for PCA. A meaningful PCA has to take into account potential high curvature; obviously, any method relying on a Euclidean linearization of tangent space will not perform well in such cases. In this paper we develop the notion of *geodesic principal components* (GPCs) based on the intrinsic distance of data to geodesics that reflect the manifold curvature. For quotient spaces this requires the notion of generalized geodesics.

Shape Spaces. In many statistical applications, data on a sub-manifold of Euclidean or Hilbert space are considered up to an isometric smooth Lie group action. Very prominently, this is the situation in the field of statistical shape analysis: *similarity shapes* are defined modulo the group of similarity transformations, e.g. Bookstein (1978) and Kendall (1984), *affine shapes* modulo the affine group, e.g. Ambartzumian ((1990), Chapter 4), and *projective shapes* modulo the general projective group, e.g. Goodall and Mardia (1999), as well as Mardia and Patrangenaru (2001). More precisely, in order to study the *shape* of a geometrical object, either a finite number of landmarks at specific locations or a bounding contour or surface is extracted and mapped to a point in a suitable Euclidean or Hilbert space. When considering similarity shapes, usually size and location information is removed by mapping onto a unit-sphere called *pre-shape-space*. Then, rotation information is removed by further mapping to elements of the quotient of the pre-shape space modulo an orthogonal group action, cf. Section 5.1. An overview of many newly developed shape space models can be found in Krim and Yezzi (2006). Earlier, finite dimensional, landmark-based shape spaces have been extensively studied; we mention the monographs of Small (1996), Dryden and Mardia (1998), as well as Kendall et al. (1999).

GPCA in Shape Analysis. One main field of application of statistical shape analysis is the study of shapes of biological entities. Methods for Kendall's landmark-based similarity shape spaces have led Le and Kume (2000) to the belief that

“biological shapes evolve mainly along geodesics”.

In joint research with the Institute for Forest Biometry and Informatics at the University of Göttingen studying the growth of individual tree stems, the biological geodesic-hypothesis is of high interest. As we will see in Section 6.1, Euclidean PCA is not applicable since the shapes in question come to lie in a region of Kendall’s shape space with unbounded curvature. This is due to the fact that shapes of tree stems are roughly degenerate long straight line segments that are invariant under rotations orthogonal to the stem, and thus correspond to singularities of the space. Clearly, many more objects in biological research such as protein structures and cell filaments are nearly one-dimensional, whereas their shape change extends into all three spatial directions. All such shapes fall into high curvature regions of shape space rendering the current methods of Euclidean PCA unsuitable.

In addition to high curvature effects, oscillation of geodesics may cause Euclidean approximations to deviate considerably from the respective intrinsic quantities. This effect is illustrated by an example in Section 6.2 of less concentrated data in regions of lesser curvature.

It is the objective of this work to propose *geodesic principal component analysis* (GPCA) that is dependent on the intrinsic structure only, and independent of a specific linearization due to an embedding into or projection onto Euclidean space. This can then be used in general on quotient spaces

1. to carry out PCA in high curvature regions and near locations where the quotient space ceases to carry a natural manifold structure,
2. to include the effects of oscillation of geodesics for less concentrated data in PCA, and
3. as a tool for detection of curvature within a data sample.

This task faces several challenges from differential geometry, statistical theory, and numerical optimization. In this work, we introduce some key concepts and major results. Many issues are beyond the scope of this paper and leave room for further research and for discussion in Section 7. This extends in particular to numerical performance and convergence issues of the algorithms employed. In our implementation, we have used standard numerical methods to locate minima. Further research is certainly necessary to derive specific fast-converging algorithms.

Throughout this paper we assume that a random variable is given on a quotient space $Q = M/G$ that arises from the isometric action of a Lie group G on a complete Riemannian manifold M . Since we know not of any application involving non-Hausdorff quotients due to non-proper actions of groups, we assume that G is compact (which is in fact somewhat more restrictive than a proper action, cf. Section 2.2). Then the quotient carries a natural metric structure, it is even locally a manifold away from *singular* locations.

The outline of this paper is as follows. In the next section we provide some background from differential geometry: at every $p \in M$ the tangent space decomposes into a horizontal and a vertical subspace. In fact, for geodesics on M , horizontality at one point is equivalent to horizontality at all points. Calling projections of horizontal geodesics on M *generalized*

Term	Description
extrinsic mean (EM)	Procrustes mean or, projection of the Euclidean mean for manifolds
intrinsic mean (IM)	minimizer of expected squared intrinsic distance
Euclidean PCA (EPCA)	based on the empirical covariance matrix in a tangent space
general Procrustes analysis (GPA)	EPCA of data projected to the tangent space at an EM
principal geodesic analysis (PGA)	EPCA of data mapped under the inverse Riemann exponential to the tangent space at the IM
geodesic PCA (GPCA)	PCA based on minimization of intrinsic residual distances to geodesics
geodesic principal components (GPCs)	minimizing geodesics of GPCA
principal component mean (PM)	intersection point of first and second GPC
restricted GPCA	GPCA while requiring that all GPCs pass through the IM
manifold PCA (MPCA)	PCA based on non-nested submanifolds totally geodesic (at a point) of increasing dimension, determined by minimizing intrinsic residual distances

Table 1: *Terminology and description of various approaches of PCA for non-Euclidean data. For the Euclidean methods on quotients such as Procrustes analysis, usually the tangent space of the original space with data optimally positioned w.r.t. the mean is used. For the intrinsic methods on quotients, we use generalized geodesics and submanifolds.*

geodesics on Q (as in Kendall et al. ((1999), pp.109–113) for Kendall’s shape spaces) we obtain a family of curves qualifying for principal components. Orthogonal projection - the corner stone of GPCA - will be defined by lifting to the manifold. In Appendix A we show that focal points, these are points with multivalent projection, and foci form a null set on the quotient. Hastie and Stuetzle ((1989), p.515) had proved this fact for one-dimensional, and Bhattacharya and Patrangenaru ((2003), p.12) for arbitrary submanifolds of Euclidean space. Thus, geodesic projections on generalized geodesics are uniquely determined up to a set of measure zero. We note that *medial axes*, introduced early to shape analysis by Blum and Nagel (1978), and currently of high interest in computer vision and in shape representation, e.g. Pizer et al. (2003), as well as Fuchs and Scherzer (2007), are taken from foci and focal points. This section is concluded by pointing to the possible oscillating and not-everywhere minimizing nature of geodesics.

In Section 3 we elaborate on basic statistical quantities on quotients as above. Unlike for Euclidean geometry in which means, variance and principal components have several equivalent

characterizations that allow for an explicit computation, in general each characterization leads to an essentially different generalization on the quotient, which in turn leads to an optimization problem that can only be solved numerically. We motivate our definition based on the minimization of residual distances. Close inspection shows that the first geodesic principal component (GPC), defined as the geodesic approximating the data best, may no longer pass through the IM, cf. Huckemann and Ziezold (2006). This fact leads to a third generalization of a mean which we call a *principal component mean* (PM); it will play a crucial role in the sections to come.

In Section 4 we lay out how to obtain sample GPCs for general quotients by pulling the numerical computation back to the manifold M . The algorithmic ansatz based on Lagrange-minimization is twofold: first computing the quotient-space distance to horizontal geodesics, thus determining *optimally positioned* data points, and second, finding all *horizontal* directions at a given data point and choosing a suitable iterate.

In Section 5 an implementable algorithm for Kendall's shape spaces is provided. Along the way we give a new and constructive proof for the fact that every singular shape can be approached by a geodesic along which some sectional curvatures are unbounded. Also, we further discuss oscillating and not-everywhere minimizing geodesics.

In the penultimate section we illustrate the effects of unbounded curvature and oscillating, not-everywhere minimizing geodesics with some exemplary 3D datasets. High curvature is encountered in the previously introduced dataset of tree stems. We find near singular shapes where

1. approximating the IM by the EM is fairly accurate; however,
2. Euclidean PCA fails to catch essential features of the shape distribution that appear under GPCA.

A classical dataset of iron age fibulae from Hodson, Sneath, and Doran (1966) serves as an example for oscillating and not-everywhere minimizing geodesics in lower curvature regions. As above, for this dataset the EM and the IM are rather close to one another. Due to oscillation, however, Euclidean PCA again fails to recognize essential features only found by geodesic PCA. This gives new results characterizing the temporal evolution of shape of these iron age brooches.

Only when both ambient curvature is low and data concentration is high, as is demonstrated by a third data set of macaque skulls, is the Euclidean approximation valid.

We note that Fletcher et al. (2004) have also proposed principal component analysis for manifolds based on geodesics. However, they require GPCs to pass through the IM and compute them by Euclidean approximation only. With our method of GPCA, the restricted GPCs through the IM can be computed as well. For applications in high curvature regions this constraint makes the restricted method as unsuitable as the Euclidean approximation. It is the additional effort to determine the location of the PM that is considerably far from the EM and IM that is crucial to the success of our method of GPCA in such cases.

2 Lie Group Action, Horizontal Geodesics and Optimal Positioning

In this section we collect well-known facts from Riemannian geometry, (e.g. Abraham and Marsden (1978), Bredon (1972), Helgason (1962), and Lang (1999)) and simple consequences thereof that are necessary to introduce notation, formulate and build up our method of GPCA. We give a comprehensive introduction not found elsewhere, as these results are not easily accessible for statisticians.

Throughout this paper we consider a connected Riemannian manifold M and a Lie group G , with Lie algebra \mathfrak{g} and unit element e , acting smoothly on M . The action will be denoted by $p \xrightarrow{g} gp$ for $p \in M, g \in G$. We also assume for the entire paper that the action is *effective*, i.e., that for every $g \neq e$ there is a $p = p_g \in M$ with $gp \neq p$. As usual $d_M(\cdot, \cdot)$ denotes the distance on M induced by the Riemannian metric.

We remark in advance that in many recent shape space models (e.g. Krim and Yezzi (2006)) infinite-dimensional Hilbert manifolds are considered. These are limits of finite-dimensional manifolds on which numerical computations are carried out. Even though many of the following results are also true in the general case of an infinite-dimensional Banach Lie group acting on an infinite-dimensional Riemannian Hilbert manifold, we note that a cornerstone of our efforts, the existence of geodesics of minimal length, Section 2.1 below, is false in general; a counterexample can be found in Lang ((1999), pp.226–227). In the following we mention explicitly if a result holds only for finite-dimensional manifolds.

2.1 Riemannian Metric and Projection to the Quotient

Denote by $\Gamma(M)$ the space of all maximal (w.r.t. inclusion) geodesics on M . The *Hopf-Rinow Theorem* asserts that on a complete Riemannian manifold geodesics $t \rightarrow \gamma(t)$ are defined for all $t \in \mathbb{R}$. Also, if M is finite-dimensional, any two points p_1, p_2 can be joined by a geodesic of length $d(p_1, p_2)$.

The Riemannian metric is denoted as usual by $p \mapsto \langle Z_p, W_p \rangle$, $p \in M$ for $Z, W \in T(M)$. Here $T(M)$ is the module of smooth vector fields on M , and $Z_p \in T_p M$ is the value of Z in the tangent space $T_p M$ of M at $p \in M$. $dg : T_p M \rightarrow T_{gp} M$, which denotes the differential induced by the action, is an isomorphism. The action of G is called *isometric* if

$$\langle Z_p, W_p \rangle = \langle (dgZ)_{gp}, (dgW)_{gp} \rangle \quad \forall p \in M, g \in G, Z, W \in T(M).$$

Then,

$$\gamma \text{ geodesic} \Leftrightarrow g\gamma \text{ geodesic} \quad \forall g \in G.$$

For $p \in M$ let $[p] = \{gp : g \in G\}$ be the *fiber* (or *orbit*) of p , and let $I_p = \{g \in G : gp = p\}$ the *isotropy* group at p . Then $[p]$ is a sub-manifold of M (locally an embedding, but in general not globally) that is diffeomorphic to G/I_p . G is said to be acting *freely* on M if all isotropy groups consist of the unit element only, i.e. $[p] \cong G \quad \forall p \in M$.

The tangent space $T_p M$ of M at p decomposes into a *vertical subspace* $T_p[p]$, that is the tangent space of the fiber, and an orthogonal *horizontal subspace* $H_p M$,

$$T_p M = T_p[p] \oplus H_p M.$$

A curve $t \mapsto \gamma(t)$ on M is called *horizontal* (*vertical*) at t_0 if its derivative there is horizontal (vertical), i.e., $\dot{\gamma}(t_0) \in H_{\gamma(t_0)} M$ ($\dot{\gamma}(t_0) \in T_{\gamma(t_0)}[\gamma(t_0)]$). Denote by $\Gamma^H(M)$ the space of all geodesics that are horizontal everywhere.

The Riemann exponential \exp_p maps a sufficiently small tangent vector $v \in T_p M$ to the point $\gamma_{p,v}(1) \in M$ when $\gamma_{p,v}$ is the geodesic through $p = \gamma_{p,v}(0)$ with initial velocity $v = \dot{\gamma}_{p,v}(0)$, i.e.,

$$\exp_p(tv) := \gamma_{p,v}(t).$$

Every point p_0 on a Riemannian manifold has a *normal neighborhood* U , i.e., for all $p \in U$ $\exists r_p > 0$ such that the inverse exponential $\log_p := (\exp_p)^{-1}$ is well defined on the *geodesic ball* $B_{r_p}(p) := \exp_p(\{v \in T_p M : \|v\| < r_p\})$ and $U \subset B_{r_p}(x)$. The *Gauss Lemma* asserts that \exp_p -images of spheres in $T_p M$ are orthogonal to geodesics through p .

Let

$$\pi: M \rightarrow M/G := \{[p] : p \in M\} \quad (1)$$

be the *canonical projection* to the quotient space equipped with the quotient topology. Note that π is both open and continuous. Then

$$d_{M/G}([p_1], [p_2]) := \inf_{g, h \in G} d_M(gp_1, hp_2) \quad \forall [p_1], [p_2] \in M/G$$

is a quasi-metric on M/G . In case of an isometric action we have that any geodesic segment γ joining p_1 and p_2 has the same length as the geodesic segment $g\gamma$ joining gp_1 and gp_2 . Hence in case of an isometric action, $d_M(gp_1, gp_2) = d_M(p_1, p_2) \quad \forall p_1, p_2 \in M, g \in G$, and thus

$$\inf_{g \in G} d_M(gp_1, gp_2) = d_{M/G}([p_1], [p_2]) \quad \forall p_1, p_2 \in M.$$

2.2 The Slice Theorem and Killing Vector Fields

In all applications we know of, M/G is a Hausdorff space which means that all fibers $[p]$ are closed in M . This is the case if G acts *properly* on M , i.e., if for all $p_n, p, p' \in M, g_n \in G, n \in \mathbb{N}$ with $g_n p_n \rightarrow p', p_n \rightarrow p$:

$$g_n \text{ has a point of accumulation } g \in G \text{ with } gp = p'.$$

A sufficient condition for a proper action is that G is compact. Even if M/G is Hausdorff the dimensions of the fibers may vary along M . Then, M/G will fail to have a natural manifold structure. This is the case for Kendall's shape spaces of three and higher-dimensional configurations. In case of a Lie group G acting isometrically and properly on a finite-dimensional manifold M , Mostov's *Slice Theorem*, cf. Palais ((1960), p.108) and Palais (1961), asserts that

for an open disk D about the origin in H_p , the twisted product $G \times_{I_p} D$ is diffeomorphic to a tubular neighborhood of $[p]$ in M . As a consequence,

$$I_{p'} \quad \text{is a subgroup of } I_p \quad \text{for } p' \in \exp_p(D). \quad (2)$$

Hence in case of a free action, $H_p M$ is locally diffeomorphic to M/G at $[p]$. Then, M/G has a unique manifold structure compatible with its quotient topology (Abraham and Marsden ((1978), p.266)), making the projection (1) a *Riemannian submersion*. Moreover then, any vector field $Z \in T(M/G)$ has a unique *horizontal lift* $\tilde{Z} \in T(M)$, i.e., $\tilde{Z}_p \in H_p M \quad \forall p \in M$. Also, every smooth curve $t \rightarrow \gamma(t)$ on M/G through $\gamma(t_0) = [p]$ has a unique horizontal lift $t \rightarrow \tilde{\gamma}(t)$ through p .

If G is compact, then any inner product on \mathfrak{g} can be extended to a bi-invariant Riemannian metric on G making all the curves $t \mapsto \text{Exp}(tv)$ geodesics on G that are defined for all $t \in \mathbb{R}$. Here $v \in \mathfrak{g}$ is arbitrary and $\text{Exp} : \mathfrak{g} \rightarrow G$ denotes the *Lie-exponential*. In most applications G is a transformation group and \mathfrak{g} is equipped with the standard Euclidean inner product. The Lie-exponential is then the exponential function for matrices. Note that $t \rightarrow \text{Exp}(tv)p$ is usually not geodesic on M .

The action of G on M gives rise to a natural mapping $\alpha : \mathfrak{g} \rightarrow T(M)$ defined by the homomorphism:

$$\begin{aligned} \alpha_p &: \mathfrak{g} \rightarrow T_p[p] \\ v &\mapsto \left. \frac{d}{dt} \right|_{t=0} \left(\text{Exp}(tv)p \right). \end{aligned}$$

In case of an isometric action, every $\alpha(v)$ is a *Killing vector field* on M . Since the local flow of a Killing vector field X is isometric, one can show

$$\frac{d}{dt} \langle X_{\gamma(t)}, \dot{\gamma}(t) \rangle = 0 \quad (3)$$

for all geodesics $t \mapsto \gamma(t)$.

2.3 Generalized Geodesics and Optimal Positioning

As an immediate consequence of (3) we have

Theorem 2.1. *Let M be a Riemannian manifold and G be a Lie group acting isometrically on M . Then a geodesic on M that is horizontal at one point is horizontal at all points.*

Due to the fact that Killing vector fields are in general not of constant modulus, a similar statement for vertical geodesics is not true. In fact, for Kendall's shape spaces of configurations of dimension $m \geq 3$, there are geodesics that are vertical at isolated points only (cf. Example 5.1).

As done in Kendall et al. ((1999), pp.109–113) for Kendall's shape spaces, the concept of geodesics can thus be pushed forward also to non-manifold quotients:

Definition 2.2. Given a quotient $\pi : M \rightarrow M/G =: Q$, where M is a Riemannian manifold and G a Lie group acting isometrically on M , call a curve δ on Q a generalized geodesic on Q if it is the projection of a horizontal geodesic on M .

$$\Gamma(Q) := \{\delta = \pi \circ \gamma : \gamma \in \Gamma^H(M)\}$$

is the space of generalized geodesics on Q . For $\pi \circ \gamma$ we also write $[\gamma]$.

Generalized geodesics can be lifted to horizontal geodesics just as in the submersion case: for $\delta \in \Gamma(Q)$ with $\delta(0) = q$ there is, given $p \in q$, a unique lift $\gamma \in \Gamma^H(M)$ such that $\pi \circ \gamma = \delta$ and $\gamma(0) = p$. If $\eta \in \Gamma^H(M)$ is any other lift of δ , then $\exists g \in G$ such that $\eta(\cdot) = g\gamma(\cdot)$. Two generalized geodesics through a common point are *orthogonal there* if their lifts through one (and thus any) common point are orthogonal there.

Definition 2.3 (Ziezold (1977)). Given a manifold M and a Lie group G acting isometrically on M , points $p_1, p_2 \in M$, and $g \in G$, call the point gp_1 in optimal position to p_2 if

$$d_M(gp_1, p_2) = d_{M/G}([p_1], [p_2]).$$

Also, gp is said to be in optimal position to a curve γ on M if

$$d_M(gp, \gamma) = d_{M/G}([p], \pi \circ \gamma).$$

If G is compact then any point can be brought into optimal position to a given point and a curve, respectively. Moreover, if gp_1 is in optimal position to p_2 then $g^{-1}p_2$ is in optimal position to p_1 , and gp_1 and p_2 are called *registered*. Note that in general, optimally positioned points will not be uniquely determined. Moreover, the relation *being in optimal position* may not be transitive, see Ziezold ((1977), p.602).

Theorem 2.4. Let M be a Riemannian manifold and G a Lie group acting isometrically on M . Then any geodesic joining two points in optimal position is horizontal.

Proof. Suppose that there is a geodesic $t \mapsto \gamma(t)$ joining $p = \gamma(0)$ and $p' = \gamma(1)$ in optimal position to each other. Then, \log_p can be defined in a neighborhood U containing $\{\gamma(t) : 0 \leq t \leq 1\}$. Moreover, let $s \mapsto \delta(s)$ be any smooth curve in $[p']$ through $p' = \delta(0)$. Then, the image of δ under \log_p is a curve outside $\log_p(U) \cap \{v \in T_p M : \|v\| < d(p, p')\}$ touching at $\log_p(p')$. By the Gauss Lemma, cf. Section 2.1, the image curve of δ is hence orthogonal to the straight line through 0 and $\log_p(p')$ which is the image of γ . As δ was arbitrary, γ is thus horizontal at p' , and by Theorem 2.1 it is a horizontal geodesic. \square

The converse, that any two points on a horizontal geodesic segment are in optimal position, is not even true in general for arbitrarily close points, cf. Theorem 5.4 (b).

As a consequence of the Hopf-Rinow Theorem (Section 2.1) and Theorem 2.4 we have the following.

Corollary 2.5. Let M be a finite dimensional complete Riemannian manifold, and G a compact Lie group acting isometrically on M and $Q = M/G$. Then any two $q_1, q_2 \in Q$ are joined by a generalized geodesic of length $d_Q(q_1, q_2)$.

2.4 Orthogonal Projection and Principal Orbit Theorem

On the quotient $Q = M/G$ of a complete Riemannian manifold we can thus define *orthogonal projection*: an *orthogonal projection* q_δ of $q \in Q$ onto $\delta \in \Gamma(Q)$ is the fiber $[p'_\gamma]$ of an orthogonal projection p'_γ of p' onto γ . Here $\gamma \in \Gamma^H(M)$ is an arbitrary lift of δ , $p \in q$, and $p' = gp$ is p put into optimal position with respect to γ . The orthogonal projection may be multivalued at some points (e.g. on a sphere, when projecting a pole to the equator); these form a set of measure zero as we shall see. Recall that a subset $A \subset M$ of a finite dimensional manifold has *zero measure* in M if for every local chart (u, U) of M the set $u(U \cap A)$ has Lebesgue measure zero. A set $B \subset Q$ has measure zero in Q if its lift $\pi^{-1}(B) \subset M$ has measure zero in M .

The following theorem is a consequence of Lemma A.2 and Theorem A.5 which is stated and proven in the Appendix.

Theorem 2.6. *Let G be a compact Lie group acting isometrically on a finite-dimensional Riemannian manifold M . Given a generalized geodesic δ on Q then the orthogonal projection q_δ is unique for all $q \in Q$ up to a set of measure zero.*

Call $M^* := \{p \in M : I_p = \{id\}\}$ the *regular space* (w.r.t. the quotient $M/G = Q$) and $M^\circ := \{p \in M : I_p \neq \{id\}\}$ the *singular space*. Since we assume an effective action, $M^* \neq \emptyset$. Below, we see that M^* is a manifold, hence by Section 2.2 the projection to the *regular quotient*

$$\pi|_{M^*} : M^* \rightarrow Q^* := M^*/G$$

is a Riemannian submersion. Some sectional curvatures of Q^* may tend to infinity when approaching a singular point, cf. Theorem 5.2. The assertion (a) of the following theorem is part of the *Principal Orbit Theorem*, cf. Bredon ((1972), p.179); the assertion (b) follows from Lemma A.2 in the Appendix.

Theorem 2.7. *Let G be a compact Lie group acting isometrically and effectively on a finite-dimensional Riemannian manifold M . Then*

- (a) M^* and Q^* are open and dense in M , Q , respectively, and
- (b) any geodesic on M that meets M^* has at most isolated points in M° .

2.5 Not-Everywhere Minimizing and Oscillating Geodesics

We call a generalized geodesic $\delta \in \Gamma(Q)$

- (a) *everywhere-minimizing* if for all two points q_1, q_2 on δ the generalized geodesic segment of minimal length between q_1 and q_2 exists and is contained in δ ,
- (b) *oscillating* if there is a point $q \in Q$ such that $t \mapsto d_Q(q, \delta(t))$ has more than one strict local minimum,
- (c) *recurrent* if there is a period $\tau > 0$ such that $\delta(t + \tau) = \delta(t)$ for all $t \in \mathbb{R}$,

- (d) *asymptotic* if there is another generalized geodesic in $\Gamma(Q)$ which is approached asymptotically by δ .

Examples of Riemannian manifolds embedded in Euclidean space.

1. On a sphere, all geodesics are recurrent, non-oscillating and everywhere-minimizing.
2. On a proper ellipsoid, every meridional geodesic (from pole to pole) is not everywhere-minimizing, as two equatorial points on it are joined by the shorter equatorial geodesic.
3. In general, geodesics on a torus are oscillating and not-everywhere minimizing. Infinitely many oscillating geodesics are recurrent and infinitely many are dense.
4. On more complicated manifolds, say surfaces of revolution generated by a function with zero curvature at a critical point, there are non-recurrent asymptotic geodesics approaching equatorial geodesics, cf. e.g. Borzellino et al. (2007).

If geodesics on M are not too ill-behaved then so are generalized geodesics on the quotient.

Remark 2.8. *If all geodesics on M are recurrent, then all generalized geodesics on $Q = M/G$ are recurrent.*

Projections of everywhere-minimizing geodesics, however, may lose this property near singularities, cf. Theorem 5.4.

3 PCA Based on Generalized Geodesics for Quotients Arising from Isometric Lie Group Actions

We first ponder different approaches to principal components on a quotient space. Then, having motivated our selection, detailed specific definitions follow.

3.1 Generalizations of PCs to Non-Euclidean Spaces

In a Euclidean space, principal components can be equivalently defined by minimizing the variance of the residuals or by maximizing the variance of the projections. Also, PCs are *nested* in the following sense: given a distribution in \mathbb{R}^m , the s -dimensional linear subspace approximating the distribution best (by minimizing sum of squared distances) is the linear space spanned by the first s principal components. We note that the mean, which is the zero-dimensional subspace approximating best, can only be found by minimizing residuals.

In a non-Euclidean space, parametric submanifolds qualify naturally as candidates for principal components. For one-dimensional components, geodesics come into mind. Higher dimensional components would then be sub-manifolds spanned by geodesics, totally geodesic at a point, as proposed by Fletcher and Joshi (2007) and computed in approximation: eigenspaces of the respective covariance matrix in the tangent space at the IM are mapped to geodesics and submanifolds totally geodesic at the IM under the Riemann exponential.

Let us now inspect which building blocks of Euclidean PCA generalize to non-Euclidean spaces in a numerically feasible manner.

Since the straight line minimizing residual variance (the sum of squared distances) is uniquely determined in Euclidean geometry, save for special cases, we expect in a non-Euclidean geometry, also “some” uniqueness of (generalized) geodesics defined by minimizing residual variance. Nestedness of PCs based on residuals, however, cannot be expected, as the IM will in general no longer come to lie on a first (generalized) principal component geodesic, cf. Huckemann and Ziezold (2006).

Alternatively, let us consider straight lines and geodesics, respectively, maximizing the sum of squared distances of projections to a variable offset. In Euclidean geometry, save for special cases, such PCs and their offsets (the projection of the mean) are uniquely determined only up to a common translation orthogonal to the respective PC. In a non-Euclidean geometry, save for special cases due to curvature, we again expect uniqueness. In case of small curvature, however, we expect poor convergence properties for numerical algorithms. Indeed, experiments with data on unit-spheres, using algorithms based on maximizing projected variance derived along the lines of the algorithms developed below, often feature slow or no convergence at all. Even worse, on manifolds or quotients with recurrent geodesics (e.g. a great circle on a sphere or generalized geodesics on Kendall’s shape spaces), the desired maximum is usually local and not global. Algorithms, moreover, may converge to the global maximum attained at an offset near the antipode of the mean. To overcome this difficulty, nestedness can be required again, cf. Fletcher and Joshi (2007). To our knowledge, it is unknown whether geodesics obtained by maximizing projection variance are nested in general. Numerical experiments on spheres hint to the contrary, that the intrinsic mean does not lie on the geodesic maximizing projected variance. It would be interesting to search for an explicit example asserting this phenomenon analytically as well.

For these reasons, we consider the *minimization of the residual variance* to be the natural approach for a non-Euclidean concept of PCA. We note that in the manifold case, our approach locally gives manifold PCs totally geodesic at a point, even though we minimize residual variance w.r.t. each single spanning GPC individually and not to the whole manifold PC. The higher-order (≥ 2) manifold PCs are then nested again.

Since the concept of generalized geodesics for quotients extends naturally to generalized sub-manifolds (one possible definition is in Appendix A), it would be an interesting task to develop methods for non-nested residual higher-order manifold PCA as well.

3.2 Geodesic PCA Based on Residuals

Throughout this section, let $\pi : M \rightarrow M/G =: Q$ be the canonical projection of a complete Riemannian manifold M on which a Lie group G acts isometrically. With the induced quasi-metric $d_Q(\cdot, \cdot) = d(\cdot, \cdot)$ on Q consider, if finite,

$$E\left(d(X, q)^2\right) \quad \text{and} \quad (4)$$

$$E\left(d(X, \delta)^2\right) \quad (5)$$

for $q \in Q$, a generalized geodesic $\delta \in \Gamma(Q)$, and a Q -valued random variable X . For Kendall's shape spaces, cf. Section 5, these quantities are finite; we assume this in the following.

In applications, it is often desirable to assume that X is continuously distributed on Q with respect to the projection of Riemannian volume. If M is of finite dimension m then from any non-vanishing m -form, a Riemannian volume can be constructed. By definition, such a non-vanishing m -form exists if and only if M is orientable. If M is non-orientable, a Riemannian volume can be defined locally only. If possible then, in order to have continuity, one would assume that the support of X is contained in the projection of a subset of M which supports a non-vanishing m -form.

A point $\mu_I \in Q$ minimizing (4) is called an *intrinsic mean* (IM) of X with *total intrinsic variance*

$$V_{int}X := E\left(d(X, \mu_I)^2\right).$$

Due to positive sectional curvatures, the IM may not be uniquely determined. E.g. this is the case for a uniform distribution on a sphere. For this reason, Kendall shapes of two-dimensional triangles with i.i. standard multi-normally distributed landmarks have no mean, cf. Dryden and Mardia ((1998), p.126). In general on manifolds, non-positive sectional curvature or sufficient concentration ensure the uniqueness of the IM. In particular, on a positive sectional curvature manifold M , if the support of a distribution of a random variable Y is contained in a geodesic ball $B_r(p)$ for some $p \in M$, and if $B_{4r}(p)$ is contained in a normal neighborhood U on which positive sectional curvatures are bounded by $\kappa > 0$, then the condition $r < \frac{\pi}{2\kappa}$ ensures that Y has a unique IM, see Karcher (1977) and Le (2001).

Now again, let us consider a random variable X on Q . In view of Theorem 2.7, one might be tempted to neglect the part of the distribution of X near the singularity set $Q \setminus Q^*$, in applications. However, since sectional curvatures may be unbounded when approaching the singularity set (Section 5.2), uniqueness of intrinsic means on Q cannot be expected in general, not even for concentrated distributions.

Definition 3.1. *A generalized geodesic $\delta_1 \in \Gamma(Q)$ minimizing (5) is called a first generalized geodesic principal component (GPC) of X . A generalized geodesic $\delta_2 \in \Gamma(Q)$ that minimizes (5) over all generalized geodesics $\delta \in \Gamma(Q)$ that have at least one point in common with δ_1 and that are orthogonal to δ_1 at all points in common with δ_1 is called a second GPC of X .*

Every point μ_P that minimizes (4) over all common points q of δ_1 and δ_2 is called a principal component mean (PM). Given a first and a second GPC δ_1 and δ_2 with PM μ_P , a generalized geodesic δ_3 is a third GPC if it minimizes (5) over all generalized geodesics that meet δ_1 and δ_2 orthogonally at μ_P . Analogously, GPCs of higher order are minimizing generalized geodesics through the PM, passing orthogonally to all lower order GPCs.

One main feature of non-Euclidean geometry is the fact that in general, due to curvature, the IM will differ from the PM, cf. Huckemann and Ziezold (2006) for a detailed discussion.

Given a generalized geodesic δ of X , denote by $X^{(\delta)}$ the orthogonal projection of X onto δ . We call it the *marginal* or the *geodesic score* of X on δ . By virtue of Theorem 2.6, geodesic

scores are uniquely defined up to a null set on Q . A minimizer $\mu_I^{(\delta)}$ on δ of the function $q \mapsto E(d(X^{(\delta)}, q)^2)$ on the GPC δ will be called an *intrinsic mean of X on the generalized geodesic δ* .

In order to define variance, recall that variance in Euclidean space can be obtained equivalently by considering projections or by considering residuals each of which, in non-Euclidean geometry, yield different results, however. Suppose we are given GPCs $\delta_1, \delta_2, \dots$ with PM μ_P . Let $m \in \mathbb{N} \cup \{\infty\}$ be the dimension of M . Then, define the *geodesic variance explained by the s -th GPC*, $1 \leq s \leq m$, $s < \infty$, as *obtained by projection*

$$V_{proj}^{(s)}X := E(d(X^{(\delta_s)}, \mu_P)^2), \quad (6)$$

with total variance

$$V_{proj}X := \sum_{s=1}^m V_{proj}^{(s)}X,$$

if finite. In the finite-dimensional case $m < \infty$, we also have the *geodesic variance explained by the s -th GPC as obtained by residuals*

$$V_{res}^{(s)}X := E\left(\frac{1}{m-1} \sum_{j=1}^m d(X, \delta_j)^2 - d(X, \delta_s)^2\right),$$

with the respective total variance

$$V_{res}X := \sum_{s=1}^m V_{res}^{(s)}X.$$

Mixing both approaches yields (again for any $m \in \mathbb{N} \cup \{\infty\}$) the definition of *mixed geodesic variance*

$$\begin{aligned} V_{mix}X &:= E(d(X^{(\delta_1)}, \mu_I^{(\delta_1)})^2) + E(d(X, \delta_1)^2) \\ &= E(d(X^{(\delta_1)}, \mu_I^{(\delta_1)})^2) + E(d(X, X^{(\delta_1)})^2). \end{aligned}$$

Finally, for $m < \infty$,

$$CX := \frac{V_{proj}X - V_{res}X}{V_{int}} \quad (7)$$

can be taken for a measure of *curvature present in X* . In a Euclidean space we have $CX = 0$. On a positive sectional curvature manifold (which tends to pull geodesics together) we expect $CX \geq 0$, whereas on a negative sectional curvature manifold (which tends to push geodesics apart) we would have $CX \leq 0$. For a distribution that mainly follows a generalized geodesic we expect CX to be small even with high absolute sectional curvatures of the surrounding space.

4 Finding Sample GPCs: Computational Issues

In this section an algorithmic method to compute sample GPCs on a quotient $Q = M/G$ is proposed. We assume that the manifold M of finite or infinite dimension is implicitly defined

by

$$\begin{aligned} M &= \{x \in \mathbb{H} : \phi(x) = 0\}, \\ T_x M &= \{v \in \mathbb{H} : d\phi(x)v = 0\}, \quad x \in M, \end{aligned}$$

for a suitable smooth function $\phi : \mathbb{H} \rightarrow \mathbb{R}^n$ with $d\phi(x) : \mathbb{H} \rightarrow \mathbb{R}^n$ having full rank for all $x \in M$. Here \mathbb{H} denotes a suitable Euclidean or Hilbert space of dimension $> n$.

In landmark-based shape analysis, for example, usually \mathbb{H} is a finite-dimensional matrix space, $n = 1$, and ϕ defines a unit-hypersphere, cf. Section 5.1. For the shape space of closed planar curves of Klassen et al. (2004) using *direction functions* θ , \mathbb{H} is the space of Fourier series, $n = 3$, and

$$\phi(\theta) = \left(\int_0^{2\pi} \cos \theta(s) ds, \int_0^{2\pi} \sin \theta(s) ds, \int_0^{2\pi} \theta(s) ds \right)$$

defines a subspace of codimension 3. Here, M itself contains only shape information, the action of $G = SO(2)$ on M allows for different choices of initial points. In view of landmark-based shape analysis, this corresponds to additionally filter out cyclic relabelling of landmarks on each curve separately. For numerical feasibility, only finitely many Fourier coefficients are used.

In general, we assume that only a finite-dimensional subspace $\mathbb{H} = \mathbb{R}^d$ is considered and that $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^{d-m}$, $d > m$, yields an m -dimensional manifold M .

Further, let G be a Lie group of finite dimension l acting isometrically on M . We assume that a similar representation is possible as well:

$$G = \{g \in \mathbb{R}^f : \chi(g) = 0\}$$

for a suitable smooth function $\chi : \mathbb{R}^f \rightarrow \mathbb{R}^{f-l}$, $f - l > 0$, and $d\chi$ of full rank on G . In all applications, G will be a compact transformation group.

In shape analysis, \mathbb{H} is usually the *configuration space* (e.g. centered configurations) and M the *pre-shape space*. The function ϕ is responsible for removing size information. Sometimes, as noted above in Klassen et al. (2004), the configuration space and more aspects of shape invariance are also defined implicitly.

In our setup M is closed and thus complete. Therefore, cf. Section 2.1, maximal geodesics $t \mapsto \gamma(t)$ are defined for all $t \in \mathbb{R}$. Denote by $\langle \cdot, \cdot \rangle$ the Riemannian metric on $T_x M$ which is, in the case of an isometric embedding, the standard Euclidean inner product. By $\gamma_{x,v}$ denote the unique geodesic on M determined by $\gamma_{x,v}(0) = x, \dot{\gamma}_{x,v}(0) = v$. Here (x, v) is an element of the *tangent bundle* $TM := \cup_{x \in M} \{x\} \times T_x M$.

Furthermore, suppose that N data points $x_1, \dots, x_N \in M$ are given that project to $[x_1], \dots, [x_N] \in Q = M/G$. With the knowledge of the preceding sections, finding generalized geodesics on M/G that minimize the squared distances to $[x_1], \dots, [x_N]$ as in (5) is equivalent to finding horizontal geodesics on M that minimize squared distances to optimally positioned data points.

Thus, we have to develop two separate sets of algorithms: the first puts points into optimal position to points and horizontal geodesics; the second computes minimizing horizontal

geodesics. Therefore, our goal is the minimization of an objective function under certain constraints. In this section we derive the corresponding Lagrange equations from which fixed point equations can be obtained. Algorithms for concrete situations can be naturally derived from the latter, as illustrated in Section 5.

4.1 Optimally Positioning

Optimally Positioning with Respect to a Point. In order to bring $x \in M$ into optimal position g^*x to a given data point $y \in M$, we have to find

$$g^* = \operatorname{argmin}_{g \in G} d_M(gx, y)^2.$$

Letting $H : G \rightarrow [0, \infty) : g \rightarrow d(gx, y)^2$, we have hence to

$$\begin{aligned} &\text{find } g^* \in \mathbb{R}^f \text{ such that} \\ &H(g^*) = \inf\{H(g) : g \in \mathbb{R}^f \text{ with } \chi(g) = 0\}. \end{aligned} \tag{8}$$

A standard method to solve this non-linear extremal problem under constraints consists in employing Lagrange multipliers. Every solution $g \in \mathbb{R}^f$ of (8) also solves

$$dH + \lambda^T d\chi = 0$$

for suitable $\lambda \in \mathbb{R}^{f-l}$. In some cases, such as for Kendall's similarity shape spaces, (8) can be solved explicitly, cf. Section 5.3.

Optimally Positioning with Respect to a Geodesic. Here we are given a data point $x \in M$ and a geodesic γ on M . In order to find $g^* \in G$ placing g^*x into optimal position to γ , minimize the objective function

$$H_1(g) := d_M(gx, \gamma)^2$$

for $g \in \mathbb{R}^f$ under the constraint $\chi(g) = 0$. This will again be achieved using the method of Lagrange multipliers by solving

$$dH_1 + \lambda^T d\chi = 0$$

for $g \in \mathbb{R}^f$ and $\lambda \in \mathbb{R}^{f-l}$. Alternatively, by solving (8), a two-stage minimization is possible: for every t find

$$g(t) := \operatorname{argmin}_{g \in G} d_M(gx, \gamma(t))^2;$$

minimize

$$H_2(t) := d_M(g(t)x, \gamma(t))$$

over t in a suitable interval such that the geodesic $t \rightarrow \gamma(t)$ is traversed once.

4.2 The Vertical Space at a Given Offset

In order to determine all horizontal geodesics it is necessary to find all horizontal directions at a given offset, i.e., all directions that are orthogonal to the vertical subspace there. To this end, we explicitly determine an orthogonal base for the vertical space $T_x[x]$ at a given point $x \in M$. Recall the homomorphism $\alpha_x : \mathfrak{g} \rightarrow T_x[x]$ from Section 2.2, suppose that $l_x = \dim(G/I_x)$, and that e_1, \dots, e_l is an arbitrary but fixed base for \mathfrak{g} . In general, even in case of a free action, the image of an orthogonal base in \mathfrak{g} will no longer be orthogonal for $T_x[x]$. Hence determine an independent system $w_1 := \alpha_x^{-1}(v_1), \dots, w_{l_x} := \alpha_x^{-1}(v_{l_x})$ in \mathfrak{g} where v_1, \dots, v_{l_x} are obtained from $\alpha_x(e_1), \dots, \alpha_x(e_{l_x})$ by a Gram-Schmidt ortho-normalization:

$$\begin{aligned} v_1 &:= \frac{\alpha_x(e_{k_1})}{\|\alpha_x(e_{k_1})\|}, & \text{where } k_1 \text{ is the smallest index s.t.} \\ & & \alpha_x(e_{k_1}) \neq 0 \\ & \vdots & \\ v_{l_x} &:= \frac{\alpha_x(e_{k_{l_x}}) - \sum_{j=1}^{l_x-1} \langle \alpha_x(e_{k_{l_x}}), v_j \rangle v_j}{\|\alpha_x(e_{k_{l_x}}) - \sum_{j=1}^{l_x-1} \langle \alpha_x(e_{k_{l_x}}), v_j \rangle v_j\|} & \text{where } k_{l_x} \text{ is the smallest index s.t.} \\ & & \alpha_x(e_{k_{l_x}}) \notin \text{span}\{v_1, \dots, v_{l_x-1}\} \end{aligned}$$

The result will be denoted by the homomorphism $\beta_x : \mathfrak{g} \rightarrow T_x[x]$ defined by

$$\beta_x(e_{k_j}) = v_j = \alpha_x(w_j) \quad j = 1, \dots, l_x. \quad (9)$$

Furthermore, define the mapping $\psi : TM = \cup_{x \in M} \{x\} \times T_x M \rightarrow \mathbb{R}^{l_x}$ by

$$\psi(x, v) := \begin{pmatrix} \langle \beta_x(e_{k_1}), v \rangle \\ \vdots \\ \langle \beta_x(e_{k_{l_x}}), v \rangle \end{pmatrix}. \quad (10)$$

Then we have that a geodesic $\gamma_{x,v}$ is horizontal if and only if $\psi(x, v) = 0$.

4.3 Minimizing Horizontal Geodesics

We derive three different types of Lagrange equations, one for the first, one for the second, and one for all subsequent geodesics. In passing, we also give an equation for the intrinsic mean not involving the Riemann exponential function, as opposed to the algorithm of Le (2001). In order to compute the variance by projection (6) we also compute the intrinsic mean on a geodesic.

First Sample GPC. Define the objective function by parameterizing (5) with

$$F(x, v) := \sum_{i=1}^N d_{M/G}(\pi \circ \gamma_{x,v}, [p_i])^2 = \sum_{i=1}^N d_M(\gamma_{x,v}, g_i^* p_i)$$

for suitable $g_i^* \in G$, $i = 1, \dots, N$, placing p_i into optimal position w. r. t. $\gamma_{x,v}$.

Every unit speed horizontal geodesic $\gamma_{x,v}$ on M is uniquely determined by an offset $x \in M$, an initial direction $v \in T_x M$ of unit length, i.e., $\langle v, v \rangle = 1$, and the horizontality condition

$\psi(x, v) = 0$. Hence, define the constraining function

$$\begin{aligned} \Phi_1 : \mathbb{R}^d \times \mathbb{R}^d &\rightarrow \mathbb{R}^{2d-2m+l_x+1} \\ (x, v) &\mapsto \begin{pmatrix} \phi(x) \\ d\phi(x)v \\ \langle v, v \rangle - 1 \\ \psi(x, v) \end{pmatrix}. \end{aligned} \quad (11)$$

Finding a first GPC on M/G is thus equivalent to solving the extremal problem

$$\begin{aligned} &\text{find } (x^*, v^*) \in \mathbb{R}^d \times \mathbb{R}^d \text{ such that} \\ &F(x^*, v^*) = \inf\{F(x, v) : x, v \in \mathbb{R}^d \text{ with } \Phi_1(x, v) = 0\}. \end{aligned}$$

Again, employ a Lagrange multiplier $\lambda \in \mathbb{R}^{2d-2m+l_x+1}$ and obtain from

$$dF + \lambda^T d\Phi_1 = 0 \quad (12)$$

two fixed point equations which naturally yield an algorithm to determine the solution (x^*, v^*) , cf. Section 5.4.

Second Sample GPC and Sample PM. Given a horizontal lift $t \mapsto \gamma_{x,v}(t)$ of a first GPC, a suitable horizontal lift of a second GPC must pass through a point $y = \gamma_{x,v}(\tau)$ with an initial direction $w \in H_y M$ orthogonal to $\dot{\gamma}_{x,v}(\tau)$. Hence, with

$$\begin{aligned} \Phi_2 : \mathbb{R} \times \mathbb{R}^d &\rightarrow \mathbb{R}^{d-m+l_x+2} \\ (\tau, w) &\mapsto \begin{pmatrix} d\phi(\gamma_{x,v}(\tau))w \\ \langle \dot{\gamma}_{x,v}(\tau), w \rangle \\ \langle w, w \rangle - 1 \\ \psi(\gamma_{x,v}(\tau), w) \end{pmatrix} \end{aligned}$$

and $F_2(\tau, w) := F(\gamma_{x,v}(\tau), w)$, finding a second GPC is equivalent to solving the extremal problem

$$\begin{aligned} &\text{find } (\hat{\tau}, \hat{w}) \in \mathbb{R} \times \mathbb{R}^d \text{ such that} \\ &F_2(\hat{\tau}, \hat{w}) = \inf\{F_2(\tau, w) : \tau \in \mathbb{R}, w \in \mathbb{R}^d \text{ with } \Phi_2(\tau, w) = 0\}. \end{aligned}$$

This will again be achieved by the method of Lagrange multipliers by solving

$$dF_2 + \lambda^T d\Phi_2 = 0 \quad (13)$$

for $\tau \in \mathbb{R}$, $w \in \mathbb{R}^d$, and $\lambda \in \mathbb{R}^{d-m+l_x+2}$. For convenience, having found $\hat{\tau}$ and \hat{w} , let $v_2 := \hat{w}$ and rewrite $\gamma_{x,v}$ as $\gamma_{\hat{x}, v_1}$ where $\hat{x} := \gamma_{x,v}(\hat{\tau})$ and $v_1 := \dot{\gamma}_{x,v}(\hat{\tau})$. Note that $[\hat{x}]$ is a sample PM on Q .

Higher Order Sample GPCs. All GPCs on Q of order r , $3 \leq r \leq m$, pass through the sample PM $[\hat{x}] \in Q$, i.e., each is determined only by a horizontal initial direction $v_r \in \mathbb{R}^d$ at offset \hat{x} . In particular, v_r is perpendicular to the horizontal lifts of all lower order GPCs at \hat{x} .

Suppose now that we have already found suitable $r-1 \geq 2$ horizontal lifts $\gamma_{\hat{x}, v_1}, \dots, \gamma_{\hat{x}, v_{r-1}}$ through \hat{x} of GPCs on Q . Then, defining

$$\begin{aligned} \Phi_r &: \mathbb{R}^d \rightarrow \mathbb{R}^{d-m+l_x+r} \\ v &\mapsto \begin{pmatrix} d\phi(\hat{x})v \\ \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_{r-1} \rangle \\ \langle v, v \rangle - 1 \\ \psi(\hat{x}, v) \end{pmatrix} \end{aligned}$$

and $F_3(v) := F(\hat{x}, v)$, finding a suitable horizontal lift of a j -th GPC is equivalent to solving the extremal problem

$$\begin{aligned} &\text{find } v_r \in \mathbb{R}^d \text{ such that} \\ &F_3(v_r) = \inf\{F_3(v) : v \in \mathbb{R}^d \text{ with } \Phi_r(v) = 0\}. \end{aligned}$$

As before, this leads to the task of solving the equation

$$dF_3 + \lambda^T d\Phi_r = 0 \quad (14)$$

for $v \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}^{d-m+l_x+r}$.

Sample IM. In a similar fashion, a representative \bar{x} of an IM can be found. For this purpose consider

$$T(x) := \sum_{i=1}^N d_Q([x], [p_i])^2 = \sum_{i=1}^N d_M(x, h_i^* p_i)^2$$

with suitable $h_i^* \in G$ for $i = 1, \dots, n$, putting $h_i^* p_i$ into optimal position to x . Then, finding a representative of an IM is equivalent to solving the extremal problem

$$\begin{aligned} &\text{find } \bar{x} \in \mathbb{R}^d \text{ such that} \\ &T(\bar{x}) = \inf\{T(x) : x \in \mathbb{R}^d \text{ with } \phi(x) = 0\}. \end{aligned}$$

The method of Lagrange multipliers yields

$$dT + \lambda^T d\phi = 0$$

for $x \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}^{d-m}$.

Sample IM on a GPC. Here, given a horizontal lift $t \mapsto \gamma(t) := \gamma_{x,v}(t)$ of a GPC $\pi \circ \gamma$ on Q , we want to find a representative $\bar{x}^\gamma = \gamma(\bar{t})$ of a point $[\bar{x}^\gamma]$ on $\pi \circ \gamma$ best approximating the orthogonal projections $[q_i]$ of the data points $[p_i]$ onto $\pi \circ \gamma$ ($i = 1, \dots, N$). Bringing p_i into optimal position $g_i^* p_i$ with respect to $\gamma_{x,v}$, observe that the orthogonal projections q_i^* of $g_i^* p_i$ onto $\gamma_{x,v}$ are representatives of $[q_i]$, $i = 1, \dots, N$. Their intrinsic mean \bar{x}^γ on $\gamma_{x,v}$ is obviously a representative for the intrinsic mean of $[q_i]$ ($i = 1, \dots, N$) on $[\gamma]$. This leads to an unconstrained extremal problem for

$$T_1(t) := \sum_{i=1}^N d_M(\gamma_{x,v}(t), q_i^*)^2$$

in one variable $t \in \mathbb{R}$.

5 Application: Kendall's Shape Spaces

We will now illustrate the generic method developed by explicitly determining functions H, H_1 , etc., and suitable χ, ϕ and ψ for Kendall's shape spaces; those are defined in the beginning. Then we apply the method of PCA based on generalized geodesics and give explicit algorithms.

5.1 Kendall's Shape Space

Kendall's shape spaces are spheres in a matrix space modulo similarity transformations; in some ways they are generalizations of complex projective spaces. Denote by

$M(m, k)$	all real matrices having m rows and k columns with the Euclidean structure of \mathbb{R}^{mk} , i.e., the inner product $\langle a, b \rangle := \text{trace}(ab^T)$, $\ a\ := \sqrt{\langle a, a \rangle}$,
$\mathfrak{gl}(m)$	$:= M(m, m)$, the Lie algebra of the general linear group $GL(m)$; the Lie exponential is then simply the matrix exponential $\text{Exp}(A) = e^A$ for $A \in \mathfrak{gl}(m)$,
$O(m)$	the orthogonal group in $GL(m)$,
$\mathfrak{o}(m)$	the Lie algebra of $O(m)$, i.e., the skew symmetric matrices in $\mathfrak{gl}(m)$,
$SM(m)$	the orthogonal complement of $\mathfrak{o}(m)$, i.e., the symmetric matrices in $\mathfrak{gl}(m)$,
$SO(m)$	$:= \text{Exp}(\mathfrak{o}(m))$, the special orthogonal group in $GL(m)$ of dimension $\frac{m(m-1)}{2}$,
i_m	$:= \text{diag}(1, \dots, 1) \in SO(m)$, the identity matrix (the unit element).

Labelled landmark-based shape analysis is based on *configurations* consisting of $k \geq m + 1$ labelled vertices in \mathbb{R}^m called *landmarks* that do not all coincide. A configuration is thus a matrix with k columns, each an m -dimensional landmark vector. Disregarding center and size, these configurations are mapped to the *pre-shape space* sphere

$$S_m^k := \{x \in M(m, k-1) : \|x\| = 1\}.$$

This can be done by, say, multiplying by a sub-Helmert matrix, cf. Dryden and Mardia (1998) for a detailed discussion of this and other normalization methods. The pre-shape sphere will be equipped with the natural spherical Riemannian metric, i.e., $T_x S_m^k$ is identified with the Euclidean space $\{v \in M(m, k-1) : \langle x, v \rangle = 0\}$.

In order to filter out rotation information, define on S_m^k a smooth action of $SO(m)$ by the usual matrix multiplication $S_m^k \xrightarrow{g} S_m^k : x \mapsto gx$ for $g \in SO(m)$. The orbit $[x] = \{gx : g \in SO(m)\}$ is the *shape* of $x \in S_m^k$. The quotient

$$\pi : S_m^k \rightarrow \Sigma_m^k := S_m^k / SO(m)$$

is called Kendall's similarity *shape space*. Since $SO(m)$ is compact, this is a Hausdorff space, cf. Section 2.2. Horizontal and vertical subspace can be explicitly determined: $v \in H_x S_m^k$ if

and only if $\text{trace}(vx^T h^T) = 0 \quad \forall h \in \mathfrak{o}(m)$, i.e.,

$$v \in H_x \Leftrightarrow vx^T \in SM(m),$$

cf. Kendall et al. ((1999), p.109). The complete situation is given by

$$\begin{array}{rcc} \mathfrak{gl}(m) & = & \mathfrak{o}(m) \oplus SM(m) \\ & & \downarrow \cdot x \qquad \qquad \uparrow \cdot x^T \\ T_x S_m^k \oplus N_x S_m^k & = & T_x[x] \oplus \overbrace{H_x S_m^k \oplus N_x S_m^k} \end{array} \quad (15)$$

Here, $N_x S_m^k := \{\lambda x : \lambda \in \mathbb{R}\}$ is the *normal space* of the pre-shape sphere. The first map is surjective, and for $\text{rank}(x) \geq m-1$, i.e., $I_x = \{i_m\}$, the second map is also surjective.

Note that the differential mapping of tangent spaces $dg : T_s S_m^k \rightarrow T_{g_s} S_m^k$ is given by

$$dg v = gv, \quad v \in T_x M, \quad g \in SO(m). \quad (16)$$

Unit speed geodesics on the pre-shape sphere are precisely the great circles

$$\gamma_{x,v}(t) := x \cos t + v \sin t \quad (17)$$

through an offset $x = \gamma_{x,v}(0) \in S_m^k$ with initial velocity $v = \dot{\gamma}_{x,v}(0) \in S_m^k$, $\langle x, v \rangle = 0$. For any $p, q \in S_m^k$, the spherical distance is given by

$$\begin{aligned} 0 \leq d_{S_m^k}(p, q) &= 2 \cdot \arcsin \left(\frac{\sqrt{\langle p-q, p-q \rangle}}{2} \right) \\ &= \arccos \langle p, q \rangle = \arccos \left(\text{trace}(pq^T) \right) \leq \pi. \end{aligned}$$

The distance of a point $p \in S_m^k$ to the great circle $\gamma_{x,v}$ is given by

$$0 \leq d(p, \gamma_{x,v}) = \arccos \sqrt{\langle p, x \rangle^2 + \langle p, v \rangle^2} \leq \frac{\pi}{2}, \quad (18)$$

and the orthogonal projection of p onto $\gamma_{x,v}$ by

$$\frac{\langle x, p \rangle x + \langle v, p \rangle v}{\sqrt{\langle x, p \rangle^2 + \langle v, p \rangle^2}}.$$

Let us now return to the action of $SO(m)$ on S_m^k , cf. Kendall et al. (1999) for a detailed discussion.

In case of $m=1$ the action is trivial, i.e., $\Sigma_1^k \cong S_1^k$.

In case of $m=2$ the action of $SO(m)$ on S_2^k is just the scalar action of $SO(2) \cong S^1 \subset \mathbb{C}$ on the $(2k-3)$ -dimensional pre-shape sphere naturally embedded in complex vector space

$$S_2^k \cong S^{2k-3} \subset \mathbb{C}^{k-1}.$$

The quotient map is then the well known *Hopf fibration*, leading to complex projective space of dimension $k-2$:

$$\Sigma_2^k \cong S^{2k-3}/S^1 = P\mathbb{C}^{k-2}.$$

In case of $m \geq 3$ a pre-shape $s \in S_m^k$ with $0 < \text{rank}(s) = r < m - 1$ will be invariant under some rotation group, a non-trivial isotropy group of dimension $r - m - 1$. For this reason, cf. Section 2.2, the shape spaces Σ_m^k ($m \geq 3$) have no natural manifold structure.

Rotating a pre-shape $p' \in S_m^k$ into optimal position to a given pre-shape $p \in S_m^k$ can be accomplished via *pseudo singular value decomposition*

$$p' p'^T = u \mu v^T$$

where $u, v \in SO(m)$ and $\mu = \text{diag}(\mu_1, \dots, \mu_n)$ with $\mu_1 \geq \dots \mu_{m-1} \geq |\mu_m| \geq 0$. Then

$$g^* := v u^T$$

puts p' into optimal position $g^* p'$ to p (e.g. Kendall et al. ((1999), p.114)). We note that the rotation g^* , and thus $g^* p'$, is uniquely determined up to a set of measure zero. More precisely: only in case of p', p regular and $\mu_{n+1} + \mu_n > 0$ is the rotation g^* uniquely determined (cf. Kendall et al. ((1999), p.121)).

5.2 The Generalized Geodesics of Kendall's Shape Spaces

Recall Theorem 2.7 to note that $SO(m)$ acts freely on the open *regular pre-shape sphere*

$$S_m^{*k} := \{x \in S_m^k : \text{rank}(x) \geq m - 1\},$$

which is open and dense in S_m^k , making the projection to *regular shape space*

$$S_m^{*k} \rightarrow \Sigma_m^{*k} := S_m^{*k}/SO(m) \subset \Sigma_m^k$$

a Riemannian submersion. Generalized geodesics in shape space restricted to regular shape space are geodesics in the usual sense. In shape space, a generalized geodesic through a regular shape $[x] \in \Sigma_m^{*k}$ is either a single geodesic in Σ_m^{*k} , or the union of geodesics in Σ_m^{*k} and isolated singular shapes in $\Sigma_m^k \setminus \Sigma_m^{*k}$ (cf. Lemma A.2. in the Appendix). For planar shape spaces Σ_2^k , the fibers of S_2^k are spanned by single vertical geodesics. In general this is not the case.

Example 5.1. For $3 \leq m < k$, a geodesic may be vertical only at an isolated point. Consider

$$x = \frac{1}{\sqrt{m}} (i_m | 0), \quad v = \frac{1}{\sqrt{2}} (w | 0) \in S_m^k \quad \text{with } w = \left(\begin{array}{cc|c} 0 & -1 & 0 \\ 1 & 0 & 0 \\ \hline 0 & & 0 \end{array} \right) \in \mathfrak{o}(m),$$

and the geodesic $t \mapsto \gamma_{x,v}(t) = x \cos t + v \sin t$. By (15), this is vertical if and only if $\exists w_t \in \mathfrak{o}(m)$ with $v \cos t - x \sin t = w_t(x \cos t + v \sin t)$. $w_0 = w$ yields verticality at $t = 0$. For $0 < t < \pi$ however, verticality would imply $-\sqrt{2} i_m = \sqrt{m} w_t w$, and this is impossible.

As $[x] \in \Sigma_m^{*k}$ tends to $[y] \in \Sigma_m^k \setminus \Sigma_m^{*k}$, some sectional curvatures at $[x]$ tend to infinity, see Kendall et al. ((1999), pp.149–156). We give a new, short, and constructive proof for this fact.

Theorem 5.2. *In shape space Σ_m^k , $k > m > 2$, every singular shape can be approached from regular shape space Σ_m^{*k} by a generalized geodesic along which some sectional curvatures are unbounded.*

Proof. In the situation of a Riemannian submersion $S_m^{*k} \rightarrow \Sigma_m^{*k}$ we have O'Neill's formula (Lang ((1999), p.393)) for the respective curvatures and, in particular, in case of sectional curvatures at $[x] \in \Sigma_m^{*k}$ of any two orthonormal vector fields $X, Y \in T(\Sigma_m^{*k})$,

$$\begin{aligned} \text{curv}_{\Sigma^*}(X, Y)_{[x]} &= \text{curv}_{S^*}(\tilde{X}, \tilde{Y})_x + \frac{3}{4} \sum_{1 \leq r < l \leq m} \langle V_{rl}, [\tilde{X}, \tilde{Y}]_x \rangle^2 \\ &= 1 + \frac{3}{4} \sum_{1 \leq r < l \leq m} \langle V_{rl}, [\tilde{X}, \tilde{Y}]_x \rangle^2. \end{aligned} \quad (19)$$

Here \tilde{X} and \tilde{Y} denote the horizontal lifts as in Section 2.2, $[\cdot, \cdot]$ denotes the *Lie bracket* and the V_{rl} ($1 \leq r < l \leq m$) constitute a base system for $T([x])$ orthonormal at x . For any vector fields V, G, H we have the well-known (Lang ((1999), p.126/7))

$$\begin{aligned} \langle V, [G, H] \rangle &= \omega[G, H] \\ &= G\langle V, H \rangle - H\langle V, G \rangle - 2d\omega(G, H), \end{aligned}$$

where ω is the one-form dual (w.r.t. the Riemannian structure) to V , and $d\omega$ denotes its exterior derivative. The above reduces to

$$\langle V, [G, H] \rangle = -2d\omega(G, H)$$

if V is vertical and G, H are horizontal. In view of (19), in order to prove the theorem it suffices thus to provide on S_k^{*m} for

- (a) a horizontal geodesic $x(t) = p \cos t + v \sin t$, $x(t) \in S_k^{*m}$ for $0 < t < \pi$, $p \in S_k^m \setminus S_k^{*m}$, and
- (b) unit length vector fields V, G, H such that V is vertical and G, H are horizontal along $x(t)$ such that
- (c) $\lim_{t \rightarrow 0} d\omega(G, H) \rightarrow \infty$ for the dual ω of V .

In fact it suffices to give an example for Σ_3^{*4} , as this can be embedded isometrically in all higher dimensional shape spaces, cf. Kendall et al. ((1999), p.29). For any singular shape $[p] \in \Sigma_3^4 \setminus \Sigma_3^{*4}$ all landmarks are on a single line segment, hence we pick w.l.o.g. a pre-shape representative of the form

$$p = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \alpha & \beta & \gamma \end{pmatrix},$$

with $0 \neq \gamma, \alpha^2 + \beta^2 + \gamma^2 = 1$. In the following, in order to verify the respective properties "horizontal" and "vertical" we make repeated use of the decomposition (15). First note that

$$v = \frac{1}{\sqrt{\alpha^2 + \gamma^2}} \begin{pmatrix} \gamma & 0 & -\alpha \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is a unit length horizontal vector at p , hence $x(t) = p \cos t + v \sin t$ satisfies the requirement (a). Requirement (b) is met by the unit length vertical field

$$V = \frac{1}{\sqrt{\sum_{j=1}^{k-1} (x_{1j}^2 + x_{2j}^2)}} \sum_{r=1}^{k-1} (x_{1r} \partial_{2r} - x_{2r} \partial_{1r})$$

and the two constant vector fields

$$G = a \partial_{11} + \partial_{12} + b \partial_{13}, \quad H = a \partial_{21} + \partial_{22} + b \partial_{23}, \quad a = -\frac{\beta \alpha}{\alpha^2 + \gamma^2}, \quad b = -\frac{\beta \gamma}{\alpha^2 + \gamma^2},$$

horizontal along $x(t)$. In order to verify (c), consider the exterior derivative of the dual to V :

$$\begin{aligned} d\omega &= -\frac{\sum_{j=1}^{k-1} (x_{1j} d^{1j} + x_{2j} d^{2j})}{\sqrt{\sum_{j=1}^{k-1} (x_{1j}^2 + x_{2j}^2)}} \wedge \sum_{r=1}^{k-1} (x_{1r} d^{2r} - x_{2r} d^{1r}) \\ &\quad + \frac{2 \sum_{r=1}^{k-1} (d^{1r} \wedge d^{2r})}{\sqrt{\sum_{j=1}^{k-1} (x_{1j}^2 + x_{2j}^2)}}. \end{aligned}$$

Along $x(t)$ we have

$$\begin{aligned} d\omega &= \frac{1}{(\alpha^2 + \gamma^2)^{3/2} \sin t} \left((2\alpha^2 + \gamma^2) d^{11} \wedge d^{21} + \alpha \gamma (d^{11} \wedge d^{23} + d^{13} \wedge d^{21}) \right. \\ &\quad \left. + 2(\alpha^2 + \gamma^2) d^{12} \wedge d^{22} + (2\gamma^2 + \alpha^2) d^{13} \wedge d^{23} \right), \end{aligned}$$

yielding, as required,

$$d\omega(G, H) = \frac{1}{(\alpha^2 + \gamma^2)^{3/2} \sin t} \rightarrow \infty.$$

□

We note that V as introduced above is, in case of dimension $m = 2$, the only (up to the sign) vertical unit length vector field. The exterior derivative of its dual is then simply

$$d\omega = 2 \sum_{j=1}^{k-1} d^{1j} \wedge d^{2j},$$

and hence $0 \leq |d\omega(G, H)| \leq 1$ for any unit length horizontal fields G, H . Then (19) yields the sectional curvature of the complex projective spaces: Σ_2^k has for $k = 3$ constant sectional curvature 4, whereas for $k \geq 4$, the sectional curvatures assume all values between 1 and 4. Observe that only the *complex* curvature of Σ_2^k is constantly 4 also for $k \geq 4$.

In order to determine the space of (generalized) geodesics $\Gamma(\Sigma_m^k)$, introduce for any $m, k \in \mathbb{N}$, $k > m$,

$O_2(m, k) := \{(e_1, e_2) \in M(m, k) \times M(m, k) : \langle e_i, e_j \rangle = \delta_{ij}, 1 \leq i, j \leq 2\}$, an orthonormal Stiefel manifold of dimension $2mk - 3$,

$O_2^H(m, k) := \{(e_1, e_2) \in O_2(m, k) : e_2 e_1^T \in SM(m)\}$ a sub-manifold of dimension $2mk - 3 - \frac{m(m-1)}{2}$.

We thus have surjective mappings $O_2(m, k-1) \rightarrow \Gamma(S_m^k)$, $O_2^H(m, k-1) \rightarrow \Gamma^H(S_m^k) : (x, v) \mapsto \gamma_{x,v}$. Under the action of $O(2)$ from the right, given by

$$(e_1, e_2) \begin{pmatrix} a & -b \\ \varepsilon b & \varepsilon a \end{pmatrix} = (ae_1 + \varepsilon be_2, -be_1 + \varepsilon ae_2), \quad a^2 + b^2 = 1 = \varepsilon^2,$$

pairs defining the same great circle are mapped onto each other. This action is free on both $O_2(m, k-1)$ and $O_2^H(m, k-1)$, hence we can identify $\Gamma(S_m^k)$ with the Grassmannian $G_2(m, k-1) := O_2(m, k-1)/O(2)$, and $\Gamma^H(S_m^k)$ with the sub-manifold

$$G^H(m, k-1) := O_2^H(m, k-1)/O(2)$$

of dimension $2m(k-1) - 4 - m(m-1)/2$. On $O_2^H(m, k-1)$ there is also a free action of $SO(m)$ from the left, defined component-wise,

$$g(e_1, e_2) = (ge_1, ge_2),$$

that commutes with the right action of $O(2)$. (16) and (17) imply that if $(x, v) \in O_2^H(m, k-1)$ determines a horizontal geodesic $\gamma_{x,v}$ on S_m^k projecting to a generalized geodesic δ on Σ_m^k , then the horizontal geodesic determined by $(gz, gv) \in O_2^H(m, k-1)$ for given $g \in SO(m)$ projects to the same δ . For even dimensions m , $\{id, -id\}$ is contained in every isotropy group on $G^H(m, k-1)$. Choosing suitably a regular pre-shape e_1 and a singular pre-shape e_2 , we see that on $G^H(m, k-1)$ we have an effective action, of $SO(m)/\{id, -id\}$ for m even, and of $SO(m)$ for m odd, respectively. Furthermore, the action is free for $m = 2$. Hence by Sections 2.2 and 2.4, we have the following.

Theorem 5.3. *The space of all (generalized) geodesics on Kendall's shape space Σ_m^k can be given the structure of the canonical quotient*

$$\Gamma(\Sigma_m^k) \cong G_2^H(m, k-1)/SO(m),$$

with manifold part of dimension $2m(k-1) - 4 - m(m-1)$. For $m = 2$, $\Gamma(\Sigma_2^k)$ is a manifold.

We now turn to the properties of (generalized) geodesics described in Section 2.5. Not all the pathological cases may occur on Kendall's shape spaces.

Theorem 5.4. *The following hold:*

- (a) for $m < k$, all (generalized) geodesics on Kendall's shape spaces Σ_m^k are recurrent;
- (b) any generalized geodesic $t \rightarrow \delta(t)$ on Σ_m^k , $2 < m < k$, with horizontal lift of the form $\gamma : t \rightarrow x \cos t + v \sin t$ with $x \in S_m^k \setminus S_m^{*k}$, $v \in S_m^{*k}$, and $\delta(t) = [\gamma(t)] \neq [\gamma(-t)] = \delta(-t)$, is not everywhere-minimizing in any neighborhood of $[x]$;
- (c) on Kendall's planar shape spaces Σ_2^k ($k \geq 3$), all geodesics are everywhere-minimizing as well as non-oscillating.

Proof. The first assertion (a) is a consequence of the fact that (generalized) geodesics on Σ_m^k are projections of horizontal great circles which are recurrent on the pre-shape sphere, cf. Remark 2.8.

For the second assertion (b), consider $y = x \cos t + v \sin t$ and $z = x \cos t - v \sin t$, $[y] \neq [z]$, for $t \neq 0$ arbitrary small. Then it suffices to show that z is not in optimal position w.r.t. y , since then $d_{\Sigma_m^k}([y], [z]) < 2t$ and δ is not everywhere-minimizing. W.l.o.g. we may assume that

$$x = \begin{pmatrix} \alpha_1^T \\ \vdots \\ \alpha_m^T \end{pmatrix}, \quad v = \begin{pmatrix} \beta_1^T \\ \vdots \\ \beta_m^T \end{pmatrix},$$

with $\alpha_i \neq 0$ and $\alpha_i^T \beta_j = \alpha_j^T \beta_i$ for $1 \leq i, j \leq r \leq m-2$, $\alpha_i = 0$ for $i = r+1, \dots, m$, and with $\beta_1, \dots, \beta_{m-1}$ non-vanishing. Then yz^T is symmetric with diagonal vector

$$\begin{pmatrix} \|\alpha_1\|^2 \cos^2 t - \|\beta_1\|^2 \sin^2 t \\ \vdots \\ \|\alpha_r\|^2 \cos^2 t - \|\beta_r\|^2 \sin^2 t \\ -\|\beta_{r+1}\|^2 \sin^2 t \\ \vdots \\ -\|\beta_m\|^2 \sin^2 t \end{pmatrix}.$$

By hypothesis, for all $0 < |t| < \pi$, at least one of the entries is negative, hence y and z cannot be in optimal position w.r.t. one another.

To prove assertion (c), let $t \mapsto \gamma(t) = x \cos t + v \sin t$ be a horizontal great circle on the pre-shape sphere $S_2^k \ni x, v$. In order to see that $[\gamma]$ is everywhere-minimizing, it suffices to show that $d_{\Sigma_2^k}([x], [p]) = t$ for $p = xc + vs$, $c = \cos t$, $s = \sin t$, and $t \in (0, \pi/2)$. With the notation of Huckemann and Hotz (2007), the fiber $[x]$ of x is given by the points $\alpha x + \beta ix \in S_2^k$, $\alpha^2 + \beta^2 = 1$ (ix denotes a distinct unit vector spanning the vertical space at x). Then,

$$d_{\Sigma_2^k}([x], [p]) = \min_{\alpha^2 + \beta^2 = 1} d_{S_2^k}(\alpha x + \beta ix, cx + sv) = \min_{\alpha^2 + \beta^2 = 1} \arccos(\alpha c) = t.$$

From this, and the fact that $[\gamma_{x,v}(t)] = [-\gamma_{x,v}(t)] = [\gamma_{x,v}(t + \pi)]$, we infer at once that $[\gamma]$ is of length π .

To see that the projection of γ to Σ_2^k is non-oscillating, consider for arbitrary $p \in S_2^k$,

$$\begin{aligned} & \left(\cos \left(d_{\Sigma_2^k}([p], [\gamma(t)]) \right) \right)^2 \\ &= \max_{\alpha^2 + \beta^2 = 1} \left(\cos \left(d_{S_2^k}(\alpha p + \beta ip, x \cos t + v \sin t) \right) \right)^2 \\ &= \max_{\alpha^2 + \beta^2 = 1} \left(\alpha (\langle x, p \rangle \cos t + \langle v, p \rangle \sin t) + \beta (\langle x, ip \rangle \cos t + \langle v, ip \rangle \sin t) \right)^2 \\ &= (\langle x, p \rangle \cos t + \langle v, p \rangle \sin t)^2 + (\langle x, ip \rangle \cos t + \langle v, ip \rangle \sin t)^2, \end{aligned}$$

which is either constantly zero or a π -periodic function. \square

Remark 5.5. Numerical examples (cf. also Section 6.2) for $3 \leq m < k$ show that generalized geodesics on Σ_m^k are usually oscillating and not-everywhere minimizing.

As a consequence of Theorem 5.4, intrinsic means feature some similarities to intrinsic means on a cone.

Corollary 5.6. The intrinsic sample mean $\hat{\mu}_I$ of two non-degenerate shapes $[y]$ and $[z]$, with sufficiently close mirrored locations to a degenerate shape $[x]$, is closer to $[x]$ but unequal to $[x]$:

$$0 < d_{\Sigma_m^k}(\hat{\mu}_I, [x]) < d_{\Sigma_m^k}([z], [x]) = d_{\Sigma_m^k}([y], [x]).$$

Proof. W.l.o.g we use the notation of y and z as in the proof of (b) in Theorem 5.4. Then for $0 < t = d_{\Sigma_m^k}([z], [x]) = d_{\Sigma_m^k}([y], [x])$ sufficiently small,

$$gz = \begin{pmatrix} \alpha_1^T \cos t - \beta_1^T \sin t \\ \vdots \\ \gamma_r^T \cos t - \beta_r \sin t \\ \beta_{r+1}^T \sin t \\ \vdots \\ \beta_{m-1}^T \sin t \\ \epsilon \beta_m^T \sin t \end{pmatrix}$$

with $\epsilon = (-1)^{m-r-1}$, is z brought into optimal position w.r.t. y . Since, for two configurations only, the extrinsic sample mean coincides with the intrinsic sample mean, a pre-shape $w \in \hat{\mu}_I$ is given by

$$\frac{y + gz}{\|y + gz\|} \equiv \frac{1}{\sqrt{\cos^2 t + \beta^2 \sin^2 t}} \begin{pmatrix} \alpha_1^T \cos t \\ \vdots \\ \gamma_r^T \cos t \\ \beta_{r+1}^T \sin t \\ \vdots \\ \beta_{m-1}^T \sin t \\ \frac{1+\epsilon}{2} \beta_m^T \sin t \end{pmatrix}$$

with indeed

$$0 < d_{\Sigma_m^k}(\hat{\mu}_I, [x]) = \arccos \frac{\cos t}{\sqrt{\cos^2 t + \beta^2 \sin^2 t}} < t,$$

where, by hypothesis,

$$0 < \beta^2 := \|\beta_{r+1}\|^2 + \dots + \|\beta_{m-1}\|^2 + \frac{1+\epsilon}{2} \|\beta_m\|^2 < 1.$$

□

5.3 Optimally Positioning w.r.t. Horizontal Great Circles

The maximal shape distance of any two shapes $[p], [p'] \in \Sigma_m^k$ is given by $\frac{\pi}{2}$. Incidentally this is also the maximal possible distance of a shape to a generalized geodesic as the example

$$p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

teaches: $d_{\Sigma_m^k}([p], [\gamma_{x,v}(t)]) = \frac{\pi}{2}$ is constant.

Now, given a pre-shape $p \in S_m^k$ and a unit speed horizontal geodesic $\gamma_{x,v}$ on S_m^k , we adapt the methods of Section 4.1 to obtain $g^* \in SO(m)$ putting g^*p into optimal position to $\gamma_{x,v}$. According to (18), the objective function to be minimized is given by

$$H_1(g) = \arccos \sqrt{(\text{trace}(gpx^T))^2 + (\text{trace}(gpv^T))^2}.$$

Equivalently, the simpler

$$\tilde{H}_1(g) := (\text{trace}(gpx^T))^2 + (\text{trace}(gpv^T))^2$$

will be maximized. For odd dimensions m we may equivalently maximize over $O(m)$, since $\tilde{H}_1(g) = \tilde{H}_1(-g)$ and for $g \in O(m)$ either g or $-g \in SO(m)$. Alternatively, with $g(t)$ in optimal position to $\gamma_{x,v}(t)$,

$$\tilde{H}_2(t) := \text{trace}(g(t)px^T) \cos(t) + \text{trace}(g(t)pv^T) \sin(t)$$

can be maximized over $[0, 2\pi)$ for odd m , and over $[0, \pi)$ for even m . Since

$$\tilde{H}_2(t) \leq (\cos d_S(g(t)p, \gamma_{x,v}(t)))^2 = \tilde{H}_1(g(t)),$$

we obtain at once that if g^*p is in optimal position to $\gamma_{x,v}$ then it is also in optimal position to its orthogonal projection onto the geodesic.

Theorem 5.7. *Given $p \in S_m^k$ and a horizontal geodesic $\gamma_{x,v}$ on S_m^k , let $g^* = g(t^*) \in SO(m)$ put p in optimal position g^*p to $\gamma_{x,v}$, as well as in optimal position to $\gamma_{x,v}(t^*)$ for some $t^* \in [0, 2\pi)$. Then*

$$\tan(t^*) = \frac{\text{trace}(g^*pv^T)}{\text{trace}(g^*px^T)}.$$

The corresponding algorithm alternates between orthogonal projection and pairwise optimally positioning.

Basic algorithm to put p into optimal position to a horizontal great circle $\gamma_{x,v}$:

Starting with a suitable $t^{(0)}$ compute, for $n \geq 0$,

$$g^{(n+1)} := uv^T,$$

where u and v are from a pseudo singular value decomposition of $p\gamma_{x,v}(t^{(n)})^T = u\mu v^T$, and update

$$t^{(n+1)} := \arctan \frac{\text{trace}(g^{(n+1)}pv^T)}{\text{trace}(g^{(n+1)}px^T)}.$$

In general \tilde{H}_2 has several local maxima that can be accessed by choosing different suitable starting values $t^{(0)}$. As correct optimal positioning is crucial to the validity of the algorithms for computing GPCs, diligent care has to be taken to obtain a global maximum.

In order to jump out of local maxima we propose to add another optimization step to the algorithm.

Diagonal-optimization algorithm to put p into optimal position to a horizontal great circle $\gamma_{x,v}$:

Starting with a suitable $t^{(0)}$ compute, for $n \geq 0$,

$$g^{(n+1)} := u\epsilon^*v^T$$

where, as before, u and v are from a pseudo singular value decomposition of $p\gamma_{x,v}(t^{(n)}) = u\mu v^T$ and

$$\epsilon^* \in E_m = \{\epsilon = \text{diag}(\epsilon_1, \dots, \epsilon_m) : \det(\epsilon) = 1, \epsilon_j \in \{-1, 1\} (j = 1, \dots, m)\}$$

is chosen to minimize $d_S(uev^T p, \gamma_{x,v})$. Then set

$$t^{(n+1)} := \arctan \frac{\text{trace}(g^{(n+1)}pv^T)}{\text{trace}(g^{(n+1)}px^T)}.$$

Note that there are four elements in E_3 and, in general, $\sum_{0 \leq j < m/2} \binom{m}{2j}$ elements in E_m .

Numerical experiments show that the diagonal-optimization algorithm converges to the global maximum in the majority of cases. Further research is necessary to develop a faster and more reliable method.

5.4 Algorithms for GPCA and Means for Kendall's Shape Spaces

In this section we compute *sample* GPCs and *sample* means algorithmically from a data sample. For brevity we omit the prefix “sample” in the following.

Throughout this section suppose we have N m -dimensional configurations, each with k landmarks. We map these to the pre-shapes $p_1, \dots, p_N \in S_m^k$. For most experimental situations we may assume with probability 1 that the isotropy groups in question are trivial, i.e., $I_x = \{i_m\}$ at pre-shapes x where iterations are performed, cf. Theorem 2.7. Hence, the dimension of the fiber is maximal: $l_x = l$. In the general case of varying l_x , the dimension of the vertical space will have to be computed anew for every iteration step.

Picking an arbitrary but fixed base e_1, \dots, e_l , $l = \frac{m(m-1)}{2}$, for $\mathfrak{o}(m)$ note that $\alpha_x(v) = vx$ for the mapping defined in Section 2.2. Defining β_x as in (9), obtain a base $w_1 :=$

$w_1(x), \dots, w_l := w_l(x)$ for $\mathfrak{o}(m)$ mapping to an orthogonal base w_1x, \dots, w_lx for the vertical space $T_x[x]$. With these, define the function ψ as in (10). For every offset $x = x_0$ the base w_1, \dots, w_l will be computed anew. When taking derivatives with respect to x at an offset x_0 we assume w_1, \dots, w_l to be constant (still being a base in an open neighborhood of x_0) and mapping to an orthogonal base at offset x_0 .

The First GPC. Recall (18) to observe that the objective function from (12) is given by

$$F(x, v) := \sum_{j=1}^N d_{S_m^k}(q_j, \gamma_{(x,v)}) = \sum_{j=1}^N \arccos^2 \sqrt{\langle x, q_j \rangle^2 + \langle v, q_j \rangle^2}$$

for $x, v \in M(m, k-1)$, where $q_j \in [p_j]$ are in optimal position to $\gamma_{x,v}$. The constraining function (11) can be taken as

$$\Phi_1(x, v) = \begin{pmatrix} \langle x, x \rangle - 1 \\ 2\langle x, v \rangle \\ \langle v, v \rangle - 1 \\ 2\langle w_1x, v \rangle \\ \vdots \\ 2\langle w_lx, v \rangle \end{pmatrix}.$$

Abbreviate $\zeta_j := \sqrt{\langle x, q_j \rangle^2 + \langle v, q_j \rangle^2}$, $\xi_j := \frac{\arccos \zeta_j}{\zeta_j \sqrt{1-\zeta_j^2}}$, and let $\xi := 1$ for $\zeta = 1$. We assume that the optimally positioned data q_j do not have the maximal distance $\frac{\pi}{2}$ to $\gamma_{x,v}$, which means that $\zeta_i \neq 0$. Computing (12) we have, with a Lagrange multiplier $\lambda = (\lambda_1, \dots, \lambda_{l+3})$, that

$$\begin{aligned} \sum_{i=1}^N \xi_i \langle x, q_i \rangle q_i &= \lambda_1 x + \lambda_2 v + \sum_{j=1}^l \lambda_{j+3} w_j^T v, \text{ and} \\ \sum_{i=1}^N \xi_i \langle v, q_i \rangle q_i &= \lambda_2 x + \lambda_3 v + \sum_{j=1}^l \lambda_{j+3} w_j x, \end{aligned}$$

where

$$\begin{aligned} \sum_{i=1}^N \xi_i \langle x, q_i \rangle^2 &= \lambda_1, \\ \sum_{i=1}^N \xi_i \langle x, q_i \rangle \langle v, q_i \rangle &= \lambda_2, \\ \sum_{i=1}^N \xi_i \langle v, q_i \rangle^2 &= \lambda_3, \text{ and} \\ \sum_{i=1}^N \xi_i \langle v, q_i \rangle \langle w_j x, q_i \rangle &= \lambda_{j+3} \text{ for } 1 \leq j \leq l. \end{aligned}$$

Letting

$$\begin{aligned} \Psi_1(x, v) &:= \sum_{i=1}^N \xi_i \langle x, q_i \rangle q_i - \lambda_2 v - \sum_{j=1}^l \lambda_{j+3} w_j^T v \\ \Psi_2(x, v) &:= \sum_{i=1}^N \xi_i \langle v, q_i \rangle q_i - \lambda_2 x - \sum_{j=1}^l \lambda_{j+3} w_j x \end{aligned}$$

we obtain the following.

Algorithm for (x^*, v^*) determining a first GPC:

Starting with initial values, e.g.

$$x^{(0)} := p_1, \quad v^{(0)} := \text{unit horizontal projection of } p_2 - p_1 \text{ at } x^{(0)},$$

obtain

$$x^{(n+1)}, v^{(n+1)} \quad \text{from} \quad x^{(n)}, v^{(n)} \quad \text{for } n \geq 0$$

by computing $q_j \in [p_j]$, $1 \leq j \leq N$, in optimal position with respect to $\gamma_{x^{(n)}, v^{(n)}}$ according to Section 5.3, and by setting

$$\begin{aligned} x^{(n+1)} &:= \frac{\Psi_1(x^{(n)}, v^{(n)})}{\|\Psi_1(x^{(n)}, v^{(n)})\|}, \\ v^{(n+1)} &:= \text{unit horizontal projection of } \Psi_2(x^{(n)}, v^{(n)}) \text{ at } x^{(n+1)}. \end{aligned}$$

The unit horizontal projection of v at $x \in S_m^k$ is, of course, given by

$$\frac{z}{\|z\|} \quad \text{where} \quad z := v - \langle x, v \rangle x - \sum_{j=1}^l \langle w_j(x)x, v \rangle w_j(x)x.$$

The Second GPC and PM. Given a horizontal great circle $\gamma_1 = \gamma_{x,v}$ mapping to a first GPC determined by $x, v \in S_m^k$, $\Phi_1(x, v) = 0$, suppose that $\gamma_2(t) = \gamma_{y,w}(t) = y \cos t + w \sin t$, with

$$y = y(\tau) = x \cos \tau + v \sin \tau = \gamma_1(\tau)$$

for some suitable $\tau \in \mathbb{R}$, is a horizontal great circle projecting to a second GPC. According to Section 4.3, the second GPC is then obtained by minimizing the objective function

$$F(\tau, w) := \sum_{j=1}^N d^2(q_j, \gamma_{(y,w)}) = \sum_{j=1}^N \arccos^2 \sqrt{\langle y, q_j \rangle^2 + \langle w, q_j \rangle^2}$$

over $(\tau, w) \in \mathbb{R} \times M(m, k-1)$, where $q_j \in [p_j]$, $i = 1, \dots, N$, are in optimal position to $\gamma_{y,w}$. Defining

$$z = z(\tau) = v \cos \tau - x \sin \tau = \dot{\gamma}_1(\tau)$$

and inspecting the first two rows of the constraining condition $\Phi_2(\tau, w) = 0$, observe that we may alternatively employ

$$\Phi_2(\tau, w) = \begin{pmatrix} 2\langle w, x \rangle \\ 2\langle w, v \rangle \\ \langle w, w \rangle - 1 \\ 2\langle w_1 y, w \rangle \\ \vdots \\ 2\langle w_l y, w \rangle \end{pmatrix}.$$

As before, abbreviate $\zeta_j := \sqrt{\langle y, q_j \rangle^2 + \langle w, q_j \rangle^2}$, $\xi_j := \frac{\arccos \zeta_j}{\zeta_j \sqrt{1-\zeta_j^2}}$, and compute (13) with a Lagrange multiplier $\lambda = (\lambda_1, \dots, \lambda_{l+3})$:

$$\left. \begin{aligned} \sum_{i=1}^N \xi_i \langle w, q_i \rangle q_i &= \lambda_1 x + \lambda_2 v + \lambda_3 w + \sum_{j=1}^l \lambda_{j+3} w_j y, \\ \sum_{i=1}^N \xi_i \langle y, q_i \rangle \langle z, q_i \rangle &= \sum_{j=1}^l \lambda_{j+3} \langle w_j z, w \rangle \end{aligned} \right\}. \quad (20)$$

Also abbreviating

$$G(a, b) := \sum_{i=1}^N \xi_i \langle a, q_i \rangle \langle b, q_i \rangle, \text{ and } A(a, b) := \sum_{j=1}^l \lambda_{j+3} \langle w_j a, b \rangle,$$

we have, for the Lagrange multipliers,

$$\begin{aligned} G(w, x) - \sum_{j=1}^l \lambda_{j+3} \langle w_j y, x \rangle &= \lambda_1, \\ G(w, v) - \sum_{j=1}^l \lambda_{j+3} \langle w_j y, v \rangle &= \lambda_2, \\ G(w, w) - \sum_{j=1}^l \lambda_{j+3} \langle w_j y, w \rangle &= \lambda_3, \text{ and} \\ G(w, w_j y) - \lambda_1 \langle x, w_j y \rangle - \lambda_2 \langle v, w_j y \rangle - \lambda_3 \langle w, w_j y \rangle &= \lambda_{j+3}, \quad 1 \leq j \leq l. \end{aligned}$$

It seems convenient to alter the algorithm such that for every step, τ is set to zero, i.e., x and v are updated for every step to y and z . Then $\lambda_{j+3} = G(w, w_j y)$ and we consider, corresponding to (20),

$$\begin{aligned} \Psi_1(x, v, w) &:= \sum_{i=1}^N \xi_i \langle w, q_i \rangle q_i - G(w, x) x - G(w, v) v \\ &\quad - \sum_{j=1}^l G(w_j x, w) w_j x, \\ \Psi_2(x, v, w) &:= \text{the } \tau \text{ of smallest absolute value satisfying} \\ &\quad G(x, v) \cos 2\tau + \frac{\sin 2\tau}{2} (G(v, v) - G(x, x)) \\ &= A(v, w) \cos \tau - A(x, w) \sin \tau. \end{aligned}$$

By definition, $\Psi_2(x, v, w)$ is orthogonal to x and v and it is horizontal. This leads to the following.

Algorithm for determining a second GPC:

Starting with some initial values, e.g.

$$\begin{aligned} x^{(0)} &:= x, \quad v^{(0)} := v, \quad w^{(0)} := \frac{z}{\|z\|} \text{ where } w := (p_2 - p_1) \text{ and} \\ z &:= w - \langle x^{(0)}, w \rangle x^{(0)} - \langle v^{(0)}, w \rangle v^{(0)} - \sum_{j=1}^l \langle w_j x^{(0)}, w \rangle w_j x^{(0)}, \end{aligned}$$

obtain

$$x^{(n+1)}, v^{(n+1)}, w^{(n+1)} \quad \text{from } x^{(n)}, v^{(n)}, w^{(n)} \quad \text{for } n \geq 0$$

by computing $q_j \in [p_j], 1 \leq j \leq N$, in optimal position with respect to $\gamma_{x^{(n)}, w^{(n)}}$ according to Section 5.3, and by setting

$$\begin{aligned} \tau &:= \Psi_2(x^{(n)}, v^{(n)}, w^{(n)}), \\ x^{(n+1)} &:= x^{(n)} \cos \tau + v^{(n)} \sin \tau, \\ v^{(n+1)} &:= v^{(n)} \cos \tau - x^{(n)} \sin \tau, \text{ and} \\ w^{(n+1)} &:= \frac{\Psi_1(x^{(n)}, v^{(n)}, w^{(n)})}{\|\Psi_1(x^{(n)}, v^{(n)}, w^{(n)})\|}. \end{aligned}$$

Having thus found x^* , v^* and w^* , recall from Section 4.3 that $\hat{x} := x^*$ is a representative of a PM on Σ_k^m . With $v_1 := v^*$ and $v_2 := w^*$, we have the two horizontal geodesics

$$\gamma_1 := \gamma_{\hat{x}, v_1}, \quad \gamma_2 := \gamma_{\hat{x}, v_2}$$

projecting to a first and a second GPC on Σ_k^m . For simplicity set $x := \hat{x}$.

Higher Order GPCs. Suppose that we have found horizontal great circles $\gamma_{x, v_1}, \dots, \gamma_{x, v_{r-1}}$, $3 \leq r \leq m$, mapping to GPCs on Σ_m^k . Then the Lagrange equation (14), for a horizontal great circle $\gamma_{x, v}$ projecting to a r -th order GPC on Σ_m^k , is given by

$$\sum_{i=1}^N \xi_i \langle v, q_i \rangle q_i = \lambda_0 x + \sum_{s=1}^{r-1} \lambda_s v_s + \lambda_r v + \sum_{j=1}^l \lambda_{r+3} w_j x$$

with $\zeta_i := \sqrt{\langle x, q_i \rangle^2 + \langle v, q_i \rangle^2}$, q_i in optimal position to $\gamma_{x, v}$, $\xi_i := \frac{\arccos \zeta_i}{\zeta_i \sqrt{1 - \zeta_i^2}}$, and suitable Lagrange multipliers $\lambda_0, \dots, \lambda_{r+l+1} \in \mathbb{R}$ that are computed as before. We then have the following.

Algorithm for determining an r -th order GPC, $r \geq 3$:

Starting with an initial value, e.g.

$$\begin{aligned} v^{(0)} &:= \frac{z}{\|z\|} \quad \text{where } w := (p_2 - p_1) \text{ and} \\ z &:= w - \langle x, w \rangle x - \sum_{s=1}^{r-1} \langle v_s, w \rangle v_s - \sum_{j=1}^l \langle w_j x, w \rangle w_j x, \end{aligned}$$

obtain

$$v^{(n+1)} \quad \text{from} \quad v^{(n)} \quad \text{for } n \geq 0$$

by computing $q_j \in [p_j], 1 \leq j \leq N$, in optimal position with respect to $\gamma_{x, v^{(n)}}$ according to Section 5.3, and by setting

$$\begin{aligned} z^{(n+1)} &:= \sum_{i=1}^N \xi_i^{(n)} \langle v^{(n)}, q_i \rangle q_i, \\ \lambda_0 &:= \langle z^{(n+1)}, x \rangle, \\ \lambda_s &:= \langle z^{(n+1)}, v_s \rangle, \quad 1 \leq s < r, \\ \lambda_r &:= \sum_{i=1}^N \xi_i^{(n)} \langle v^{(n)}, q_i \rangle^2, \\ \lambda_{j+r+1} &:= \langle z^{(n+1)}, w_j x \rangle, \quad 1 \leq j \leq l, \text{ and} \\ v^{(n+1)} &:= \text{sign}(\lambda_r) \frac{z^{(n+1)} - \lambda_0 x - \sum_{s=1}^{j-1} \lambda_s v_s - \sum_{j=1}^l \lambda_{j+r+1} w_j x}{\|z^{(n+1)} - \lambda_0 x - \sum_{s=1}^{j-1} \lambda_s v_s - \sum_{j=1}^l \lambda_{j+r+1} w_j x\|}. \end{aligned}$$

The IM and the IM on a GPC. The computation of an IM according to Section 4.3 can be carried out analogously. With $\zeta_j := \langle x, q_j \rangle$, $\xi_j = \frac{\arccos \zeta_j}{\sqrt{1-\zeta_j^2}}$, $\sum_{j=1}^N \xi_j \langle q_j, x \rangle = \lambda$ and $\Psi(x) := \text{sign}(\lambda) \sum_{j=1}^N \xi_j q_j$, we have the following algorithm:

$$\left. \begin{aligned} x^{(n)} &\mapsto x^{(n+1)} \\ x^{(n+1)} &= \frac{\Psi(x^{(n)})}{\|\Psi(x^{(n)})\|} \end{aligned} \right\}.$$

For every iteration, all $q_j \in [p_j]$ are rotated into optimal position to $x^{(n)}$.

Computing an IM on a GPC is equivalent to computing an IM on a horizontal great circle on a sphere with respect to projections of optimally positioned points according to Section 4.3. The corresponding algorithm can be found in Huckemann and Ziezold (2006).

6 Data Examples

In conclusion, we apply our methods to three typical data sets. The first example, from forest biometry, features concentrated and nearly degenerate shapes. In the second and third examples we consider classical data sets: regular, less concentrated shapes from an archaeological site, and regular concentrated shapes of macaque skulls. For all data sets we computed

GPCA as laid out in this paper, and compared it with

restricted GPCA by restricting the GPCs to pass through the sample IM with our algorithms accordingly simplified,

Euclidean PCA at the IM (PGA) by computing the covariance matrix of the data mapped under the inverse Riemann exponential to the tangent space at the sample IM, cf. Fletcher et al. (2004),

Euclidean PCA at the EM (GPA) by computing the covariance matrix of the data orthogonally projected to the tangent space of the Procrustes sample mean, cf. Dryden and Mardia ((1998), Chapter 5).

Again for brevity we omit the prefix “sample” in the following. All GPCs, means and variances found below are in fact *sample* GPCs, *sample* means and *sample* variances.

6.1 Nearly Degenerate Shapes

In collaboration with the Institute for Forest Biometry and Informatics at the University of Göttingen, the influence on the shape of tree stems of certain external and internal factors is studied. Of particular interest is the influence by competition with nearby trees that commences when the crowns meet. We consider here a dataset of tree stems of five Douglas fir trees collected at an experimental site in the Netherlands, cf. Gaffrey and Sloboda (2000) and Table 5 in Appendix B.

GPCA	GPC1	GPC2	GPC3	GPC4	GPC5
by projection	93.58 %	6.38 %	0.04193 %	$5.309e - 06$ %	$2.505e - 06$ %
rmssd	0.0002818	0.00183	0.002405	0.002420	0.002420
GPCA through the IM	GPC1	GPC2	GPC3	GPC4	GPC5
by projection	50.23 %	49.50 %	0.2744 %	$2.197e - 05$ %	$2.390e - 07$ %
rmssd	0.0009775	0.001077	0.003392	0.0034	0.0034
Euclidean PCA at IM	PC1	PC2	PC3	PC4	PC5
Euclidean projection	67.33 %	32.43 %	0.24 %	$5.366e - 06$ %	$4.295e - 15$ %
geodesic projection	50.98 %	48.62 %	0.4076 %	$6.041e - 05$ %	
rmssd	0.001046	0.001240	0.003388	0.0034	
Euclidean PCA at EM	PC1	PC2	PC3	PC4	PC5
Euclidean projection	67.33 %	32.43 %	0.2401 %	$5.597e - 06$ %	$1.494e - 31$ %
geodesic projection	50.98 %	48.62 %	0.4073 %	$6.157e - 05$ %	
rmssd	0.001046	0.001240	0.003388	0.0034	

Table 2: *Displaying for five shapes of Douglas fir trees the percentages of variance explained by PCA based on generalized geodesics (first box), PCA based on generalized geodesics while requiring that all GPCs pass through the intrinsic mean (second box), PCA by computing the covariance matrix in the tangent space of the IM under the inverse Riemann exponential (third box), and PCA by computing the covariance matrix of the data orthogonally projected to the tangent space of the EM (bottom box). In the first two boxes, the line labelled “by projection” gives the percentages of variance obtained by projection, cf. (6); the line labelled “rmssd” reports the square-root of the mean of the squared shape-distances of the data to the respective generalized geodesic. In the lower two boxes the values obtained by “Euclidean projection” are the percentages of eigenvalues of the respective covariance matrices under the inverse Riemann exponential at the IM (Euclidean PCA at IM), and under orthogonal projection to the tangent space at the EM (Euclidean PCA at EM). Under “geodesic projection” the variances obtained by projection (as above) to the respective generalized geodesics corresponding to the eigenvectors of the respective covariance matrices are reported. No values are reported for the ultimate eigenvectors as they point no longer into horizontal space. Similarly “rmssd” gives again the square-root of the mean of the squared shape-distances of the data to the respective generalized geodesics.*

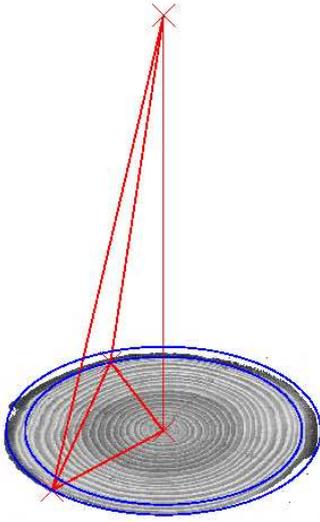


Figure 1: *Extracting a tetrahedral shape from a Douglas fir tree stem. Distorted view.*

Since crown competition will be less reflected by lower tree rings, tetrahedral shapes have been extracted by placing landmarks at the top of the tree, as well as on the center of the stem-disk at middle height, and at the maximal and minimal radius of this disk, cf. Figure 1. Since the heights in questions are about 10 meters and the radii around 10 centimeters, all shapes are close to one-dimensional line segments. Shape change, however, occurs in all three spatial directions. As is visible in Table 2, relative data variation along the respective PCs, and distances of the data to the first PC differ considerably between GPCA, restricted GPCA and Euclidean PCA. Thus, with this data-set we are very close to the situation described in Theorem 5.2: the shape space is locally spanned by vector fields (among others) whose sectional curvatures tend to infinity when approaching the singularity. Moreover, note that the difference between “variance explained by projection” and distances to generalized geodesics (corresponding to “variance explained by residuals”) can be taken as a measure for curvature present in the data.

As the data is concentrated on a geodesic (93.58% of the variation is along the first GPC), the empirical value of CX (2.466) must be generated by a high curvature of the surrounding space, cf. (7). The overall variance rmiv (the root of the total intrinsic variance divided by sample size) of the data-set is small. Still, for the same reason, the PM \bar{x}_P is considerably far away from the EM \bar{x}_E and the IM \bar{x}_I , the latter three being rather close to each other, in fact the distance between PM and IM is about of the size of rmiv . This means that the PM is well within the data spread:

$$\text{rmiv} = 0.00342, \quad d_{\Sigma_3^4}(\bar{x}_P, \bar{x}_I) = 0.002126, \quad d_{\Sigma_3^4}(\bar{x}_E, \bar{x}_I) = 2.107e - 08, \quad .$$

Therefore the classical methods of GPA, Euclidean PCA at the EM, and Euclidean PCA at the IM are practically identical. Also, when considering residuals and projections to generalized geodesics, Euclidean PCA and GPCA restricted to the IM are almost equivalent. On the other hand, due to the flexibility of GPCA in choosing the PM, the first GPC approximates the data far better than the first PCs of the other methods; in fact the approximation of the data by the first PC of restricted GPCA is worse by a factor larger than 3. Most notably, however, Euclidean PCA and the almost equivalent GPCA restricted through the IM fail to recognize the non-trivial fact that shape change for this data set is apparently one-dimensional. This finding suggests that shape variation of nearby trees of the same kind in interaction, is essentially one-dimensional, cf. Hotz et al. (2007).

	GPCA	GPC1	GPC2	GPC3	GPC4	GPC5
	by projection	29.16 %	13.48 %	48.69 %	8.152 %	0.5145 %
	rmssd	0.1178	0.1270	0.1586	0.1681	0.1720
<hr/>						
	GPCA through the IM	GPC1	GPC2	GPC3	GPC4	GPC5
	by projection	28.12 %	12.72 %	51.59 %	7.085 %	0.4866 %
	rmssd	0.1193	0.1266	0.1554	0.1685	0.1709
<hr/>						
	Euclidean PCA at IM	PC1	PC2	PC3	PC4	PC5
	Euclidean projection	48.95 %	37.96 %	10.11 %	2.236 %	0.7397 %
	geodesic projection	8.932 %	51.29 %	34.14 %	0.4577 %	5.173 %
	rmssd	0.1228	0.1241	0.1567	0.1703	0.1711
<hr/>						
	Euclidean PCA at EM	PC1	PC2	PC3	PC4	PC5
	Euclidean projection	49.53 %	37.11 %	10.32 %	2.287 %	0.7598 %
	geodesic projection	8.95 %	50.99 %	34.27 %	0.4671 %	5.326 %
	rmssd	0.1228	0.1355	0.1631	0.1704	0.1717

Table 3: Displaying for the Münsingen brooch data-set the results of the various methods of PCA with the notation of Table 2.

6.2 Non-Concentrated Regular Shapes

In a second illustration, consider a data-set of 28 brooches (fibulae) from an Iron Age grave site in Münsingen, Switzerland, closely studied by Hodson et al. (1966) among others. As the cemetery grew over a large period of time, it is reasonable to believe that brooches at different locations originated in different time epochs. Thus, five individual temporal groups have been proposed. In order to study shape change, at each of the three-dimensional brooches, four landmarks have been assigned to specific “anatomical” locations. Small ((1996), Section 3.5)

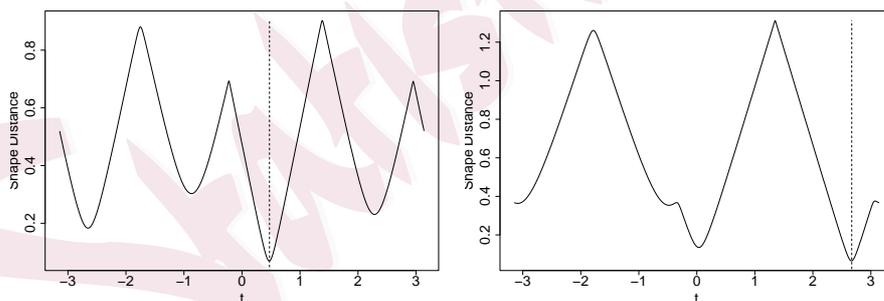


Figure 2: Shape distance (vertical axis) to shapes of the second (left) and the twenty-second (right) fibula (the sample has 28 fibulae), to shapes along the generalized geodesic $\delta(t)$, $-\pi \leq t \leq \pi$, (horizontal axis) determined by the second Euclidean PC at the EM. The dashed line marks the global minimum, its horizontal location gives the score.

has applied *principal coordinate analysis* to *Procrustes innerpoint distances* of a planar lateral view of the landmarks. Table 3 displays the relative variances explained by the respective five PCs in Kendall's shape space Σ_3^4 .

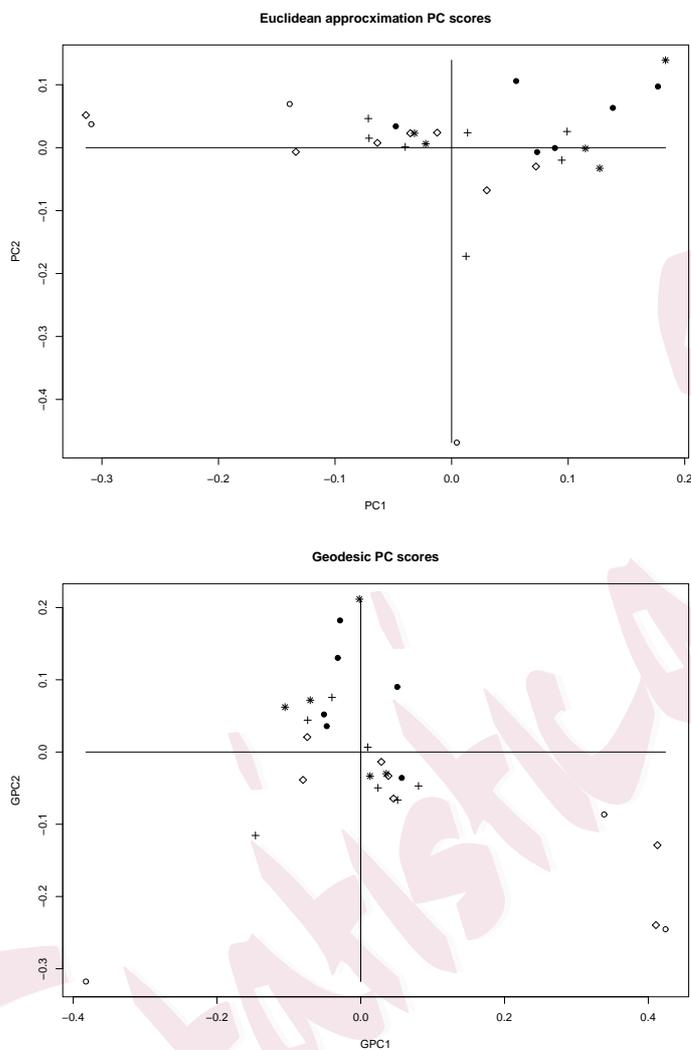


Figure 3: Scores of the five groups of the Münsingen brooch data on the first two PCs. Top: Euclidean PCA at the EM, bottom: geodesic method. The oldest group is depicted by filled circles, the second oldest by stars, the middle group by crosses, the second youngest by diamonds and the youngest by circles. The scaling on the coordinate axes is depicted in Euclidean projection (top image) and arclength on shape space (bottom image, the largest distance in shape space is $\pi/2$). Note the different scaling on the axes.

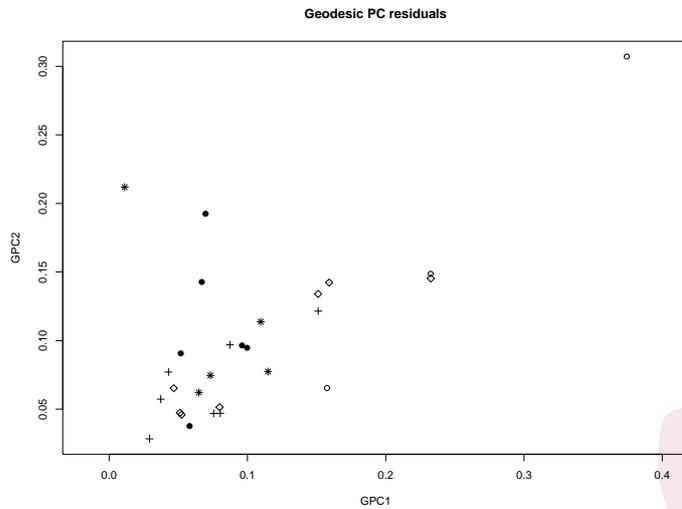


Figure 4: Residual distances to the first two GPCs for the five groups from Figure 3 of the Münsingen brooch data-set.

As visible, the results of the two geodesic methods, which are this time very similar, differ from the results of the two Euclidean approaches, which are again almost identical. Euclidean PCA leads to the belief that the variation of brooch shape is essentially explained by two PCs.

Two dominating components (minimizing residual variance) can also be found by the two geodesic methods; however, the corresponding GPCs point in directions different from the directions of the Euclidean PCs in the tangent space of the EM (or IM, resp.) and approximate the data w.r.t. the intrinsic metric better. This fact can be explained by curvature, and even more by oscillation and the not-everywhere minimizing property of the respective generalized geodesics corresponding to the Euclidean PC directions, cf. Figure 2. There, for two fibulae, oscillation of the generalized geodesic corresponding to the second Euclidean PC is visible. The left image shows a score close to zero, the right image depicts a score close to π though there is a local (slightly higher) minimum close to zero. For this reason, variance explained by projection may increase with higher order PCs for all methods, even though higher order PCs have higher residual variance. This hints again at a smaller explanatory significance of variance obtained by projection, in particular for the Euclidean approximations. Analysis of residuals will then add to the significance of the findings.

In the Euclidean approximations a trend of temporal evolution can be identified as the strongest component, see Figure 3: shapes move in time from right to left along the first PC, cf. also Small ((1996), p.94). Using the geodesic method, this trend is also visible in the second component: shapes move in time from top to bottom. As the stronger first principal component, however, a temporal increase in shape diversification can be identified. This latter observation is validated by a plot of residual distances in Figure 4.

GPCA	GPC1	GPC2	GPC3	GPC4	GPC5
by projection	31.15 %	20.10 %	14.67 %	10.57 %	6.223 %
rmssd	0.06158	0.06636	0.06858	0.07021	0.0719
GPCA through the IM	GPC1	GPC2	GPC3	GPC4	GPC5
by projection	31.15 %	20.10 %	14.67 %	10.57 %	6.223 %
rmssd	0.06158	0.06636	0.06858	0.07021	0.0719
Euclidean PCA at IM	PC1	PC2	PC3	PC4	PC5
Euclidean projection	31.18 %	20.10 %	14.67 %	10.58 %	6.216 %
geodesic projection	31.15 %	20.10 %	14.67 %	10.57 %	6.223 %
rmssd	0.06158	0.06636	0.06858	0.07021	0.0719
Euclidean PCA at EM	PC1	PC2	PC3	PC4	PC5
Euclidean projection	31.13 %	20.11 %	14.68 %	10.59 %	6.223 %
geodesic projection	31.15 %	20.10 %	14.67 %	10.57 %	6.223 %
rmssd	0.06158	0.06636	0.06858	0.07021	0.0719

Table 4: *Displaying for a data set of macaque skulls the results of the various methods of PCA with the notation of Table 2.*

In conclusion of this example, we note that the curvature estimate (7) for this data set is $CX = 0.1217$, much smaller than in the previous example. For this reason, as visible in Table 3, GPCA restricted through the IM approximates unrestricted GPCA rather well. Equivalently, in relation to the data spread (rmiv) and the again comparatively small distance between EM and IM, the PM is relatively closer to both IM and EM than in the previous example:

$$\text{rmiv} = 0.1734, \quad d_{\Sigma_3^4}(\bar{x}_P, \bar{x}_I) = 0.01936, \quad d_{\Sigma_3^4}(\bar{x}_E, \bar{x}_I) = 0.002291, \quad .$$

6.3 Concentrated Regular Shapes

In a final example we consider a data set of seven anatomical landmarks chosen on the skulls of nine male and nine female macaque specimen, cf. Dryden and Mardia (1998). Male specimen show a greater variance, the difference between mean shapes (EM) of both sexes is not significant, however, cf. Dryden and Mardia ((1998), p.159). As in the previous examples, Table 4 reports the results of the various methods of PCA. In this example we observe no difference between the three approaches, rendering the Euclidean approximation valid. In fact, the curvature estimate (7) takes the very small value $CX = 0.0001924$, and the root of the total intrinsic variance divided by sample size is also of small size, $\text{rmiv} = 0.05954$. Due to these facts, all means are nearly coincident and the various methods of PCA give (almost) identical results, cf. Table 4.

7 Discussion

With this paper, we aimed at providing a method that allows one to perform non-linear multivariate statistics in cases where the data live on spaces with a non-Euclidean intrinsic structure. Furthermore, we wanted this method to rely on the given intrinsic structure alone, thereby avoiding any linear approximation. We were particularly interested in a rather general quotient space occurring from an isometric action of a Lie group on a Riemannian manifold. For such spaces, we proposed a method as desired, namely a PCA based on the spaces' intrinsic structure. A typical application lies in the statistical analysis of shapes which has until now been performed almost exclusively by linear approximations. An approach respecting the non-linearity of the intrinsic metric leads to non-linear optimization problems, which pose specific numerical challenges. The methods derived in this work are applicable to a wide variety of non-manifold shape spaces which usually occur when studying three- and higher-dimensional shapes. Taking Kendall's shape spaces as an example we have illustrated how this methodology provides for explicit algorithms.

We have presented PCA based solely on the intrinsic structure in contrast to the current methods of PCA by embedding into and projecting to Euclidean space; our motivation was threefold. First, the intrinsic approach may be the only method available near singularities of the quotient. On Kendall's shape spaces this is the case when mean shapes are nearly singular. Recall from the remarks following Theorem 5.2 that this cannot happen with planar shapes, i.e., for Σ_2^k . It is, however, the case when samples of three-dimensional objects fall into high curvature regions of the regular shape space, as happens in biological applications. This phenomenon which we encountered in the study of shapes of individual tree stems (cf. Section 6.1) was the initial motivation for this research. Second, our approach may serve as a basis for a larger study to compare the validity of the much simpler linear approximation methods. Such a study was carried out by Huckemann and Hotz (2007) for Kendall's planar shape spaces. Third, we consider our work a preliminary basis for further development of non-linear statistical analysis working exclusively intrinsically.

At this point let us sum up our results on quotients with singularities such as Kendall's shape spaces concerning GPCA, restricted GPCA through the IM, Euclidean PCA at the IM (PGA), and Euclidean PCA at the EM (general Procrustes analysis), cf. Section 6. We found the following.

1. Due to the vicinity of IM and EM and the proximity of Euclidean and intrinsic distances, Euclidean PCA at the IM and Euclidean PCA at the EM are practically equivalent.
2. Due to unbounded curvature and oscillations of generalized geodesics, Euclidean PCA may fail to recognize data features that occur under GPCA. This is the case for
 - (a) large data-spread, even with little curvature present: then GPCA may well be approximated by restricted GPCA; and
 - (b) data within high curvature regions: then restricted GPCA may also fail to recognize

data features that occur under GPCA due to the fact that IM and PM may differ considerably when high curvature is present.

3. Curvature within the data can be estimated by CX from (7).

We view our methods and results as an early contribution to the ambitious task of carrying over statistical methods from linear Euclidean spaces to manifolds and more general spaces. We would like to conclude by pointing out open problems and research directions we consider to be of high importance.

Exploring Effects of Non-Euclidean Geometry. As we have seen, the geometric structure of generalized geodesics for Kendall's shape spaces for three-dimensional configurations is far more complicated than it is for two-dimensional configurations. In particular, the location and the descriptive nature of the PM deserves to be studied more closely. Also, studying oscillation effects certainly requires more research. In case of Kendall's shape spaces, an upper bound on the number of local minima of the distance of a given shape to shapes along a generalized geodesic depends linearly on $\dim(SO(m)/I_{\gamma(t)})$ with the isotropy group $I_{\gamma(t)}$ at $\gamma(t)$. Note that by Lemma A.2 this dimension is constant a.e. on $t \in [0, 2\pi)$. In general, conditions on geodesics and the data support D (the geodesic convex hull of the data, say) can be found that ensure that geodesics and geodesic segments in D are everywhere minimizing.

Computing Curvature Present in the Data. Even with good numerical algorithms, GPCA will be computationally more costly than Euclidean PCA. It would be helpful to develop diagnostic tools for the evaluation of the benefit of GPCA over Euclidean PCA. CX as introduced in (7) is such an indicator; however, CX is available only after computing the GPCA. An alternative could be a numerical method to compute intrinsic sectional curvature, say at a mean of the data sample, using suitable horizontal and vertical vector fields.

Degenerate Means. Sample means of non-degenerate shapes grouping around a degenerate shape may move closer to the degenerate shape, cf. Corollary 5.6. It would be interesting to determine under what circumstances (e.g. symmetry conditions on the distribution) the intrinsic population mean of concentrated non-degenerate shapes may be unique and/or degenerate.

Numerics. Crucial to the success of GPCA is a reliable method for optimal positioning w.r.t. a horizontal geodesic. The results of this paper have been validated by highly time consuming brute force methods to ensure that local minima found are in fact global. In order to undertake larger studies in the future, considerable improvement of the numerical methods is essential. The numerical problem can be attacked either by travelling along a geodesic i.e., essentially along S^1 , or by optimizing over $SO(m)$; both are problems of non-convex optimization, cf. Section 5.3. The objective function corresponding to the former can be non-differentiable, thus possibly leading to very narrow minima; the objective function corresponding to the latter can

be expressed as a multivariate polynomial. To the knowledge of the authors these problems encountered have found little attention in numerical analysis in the past.

Extensions to Other Shape Spaces. Obviously, the methods developed here extend to many other shape spaces as well. If horizontal geodesics and distances to them are given analytically, then our method can be applied directly. Our approach may be applicable as well if geodesics and distances to geodesics can only be computed numerically. Such numerical methods have been proposed, e.g. by Miller et al. (2006) or Schmidt et al. (2006), for general models, such as e.g. Michor and Mumford (2006) or Klassen et al. (2004), for closed 2D curves, or closed 2D and 3D contours based on medial axes, e.g. Pizer et al. (2003), Fletcher et al. (2004), as well as Fuchs and Scherzer (2007).

Manifold PCA. On the conceptual side, as pointed out in Section 3.1, non-nested intrinsic PCA based on (generalized, cf. Appendix A) manifolds appears as a natural extension of our methods. One might consider manifolds totally geodesic at a point as well as manifolds totally geodesic at all points. In particular for Kendall's shape spaces of three- and higher-dimensional objects, the former may be numerically hard to compute, and the latter may be available only for some dimensions. We also note that parametrized spaces obtained thereby may be manifolds only locally.

Inference. In this paper we did not address the issue of inference. Classical and functional PCA allows for manifold inferential tools, such as tests and confidence bands on the components (see e.g. Kneip and Utikal (2001) or Munk et al. (2007)). These approaches are all asymptotic in nature and a linearization of the underlying estimator typically leads to satisfactory results, often for quite small samples. This does not seem to be the case anymore in the present context, rather high curvature effects suggest that (first order) asymptotic considerations will fail. Model selection, as it is used for the automatic selection of the number of required components (see e.g. Hsieh (2007) in the context of nonlinear PCA), is another related issue. It would be of great importance to transfer these ideas to GPCA. Finally, we believe that there is much room for sharp risk bounds for the geodesic principal components as it has been investigated e.g. by Zwald et al. (2004) in the context of kernel PCA.

Appendix A: Foci and Focal Points

In this section we use the notation and hypotheses of Section 2: a compact Lie group G acts isometrically on a complete finite-dimensional and connected Riemannian manifold M giving rise to the canonical quotient $\pi : M \rightarrow M/G =: Q$. In order to define *generalized submanifolds* on Q as well as *foci and focal points*, we introduce some additional notation and results, cf. Bredon ((1972), p.182).

A point $p \in M$ is of *orbit type* (G/H) if $I_{gp} = H$ for some $g \in G$. This is equivalent to

saying that all isotropy groups in the fiber $[p]$ are conjugates of H . The union

$$M^{(H)} := \{p \in M : I_{gp} = H \text{ for some } g \in G\}$$

of all points of equal orbit type (G/H) - if not void - is a submanifold of M , and $\overline{M^{(H)}} \setminus M^{(H)}$ consists of points of *smaller orbit type*, i.e., $H \subset I_{gp'}$, $H \neq I_{gp'}$ for every $p' \in \overline{M^{(H)}} \setminus M^{(H)}$ with suitable $g = g_{p'} \in G$. In particular, $Q^{(H)} := M^{(H)}/G$ is a manifold. Recall that $Q^{\{\text{id}\}} = Q^*$ is the manifold part of Q .

It is well-known that every point has a neighborhood in which only finitely many orbit types occur. This can easily be seen by an inductive argument relying on the Slice Theorem (cf. Section 2.2) and Lemma A.2 below. Since manifolds are separable (i.e., contain a countable dense set) there are only countably many orbit types on M .

Definition A.1. We call $K \subset Q$ a generalized submanifold of Q if $K \cap Q^{(I_p)}$ is a submanifold of $Q^{(I_p)}$ for every $[p] \in K$.

Since $\pi|_{M^{(I_p)}} : M^{(I_p)} \rightarrow Q^{(I_p)}$ is a Riemannian submersion, $L^{(I_p)} := \pi^{-1}(K \cap Q^{(I_p)}) \subset M^{(I_p)}$ is a submanifold of M for all $[p] \in K$, for a generalized submanifold K . In particular, every submanifold K of Q^* is a generalized manifold. The notion of generalized submanifolds includes generalized geodesics.

Lemma A.2. The point set L of a horizontal geodesic contains a point p of maximal orbit type. Then $L^{I_p} := L \cap M^{(I_p)} = \{p' \in L : I_{p'} = I_p\}$ is a submanifold, $L \setminus L^{I_p}$ is contained in $\overline{M^{(I_p)}} \setminus M^{(I_p)}$, and consists only of isolated points.

Proof. Assume that $\gamma(t) = \exp_{p_0}(tv)$ with $p_0 = \gamma(0)$ and suitable $v \in H_{p_0}M$ is a parametrization of the point set L of a horizontal geodesic. Moreover let $H_t := I_{\gamma(t)}$. From (2) we obtain $H_t \subset H_0$ for t sufficiently small. Since

$$g(\exp_{p_0}(tv)) = \exp_{g_{p_0}}(t dg v),$$

we have at once $g \in H_t$ for $t \neq 0$ sufficiently small if and only if $dg v = v$. Hence $H_t = H_{t'}$ for all sufficiently small non-zero t, t' . This yields that the orbit type on L near p_0 is constant and maximal, except for at most the isolated point p_0 . W.l.o.g assume that $p_0 = p$ is of locally maximal orbit type. Since $\gamma(t) = \exp_p(tv)$ is a global parametrization of L , we have that $H_0 \subset H_t$ for all $t \in \mathbb{R}$. This argument can be applied to any point of locally maximal orbit type, hence $H_0 = H_t$ for all $t \in \mathbb{R}$ except for at most isolated points. \square

Now, consider an arbitrary submanifold L of M and an arbitrary generalized submanifold K of Q . Recall the normal bundle $NL = \cup_{p \in L} \{p\} \times N_p L$ of L in M . With the Riemann exponential of M we have the well defined *endpoint map*

$$\begin{aligned} \phi : NL &\rightarrow M \\ (p, v) &\mapsto \phi(p, v) = \exp_p(v). \end{aligned}$$

Also, we have the set of *orthogonal projections*

$$K_{[p]} := \left\{ [p^*] \in K : d_Q([p], [p^*]) = \inf_{[p'] \in K} d_Q([p], [p']) \right\}$$

of $[p] \in Q \setminus K$ onto K .

Definition A.3. *Call*

$$\begin{aligned} p \in M \setminus L \text{ a focus of } L & \quad \text{if} \quad \exists (p', v') \in NL \text{ such that } \exp_{p'}(v') = p \text{ and} \\ & \quad (\mathrm{d}\phi)_{(p', v')} \text{ is singular,} \\ [p] \in Q \setminus K \text{ a focal point of } K & \quad \text{if} \quad |K_{[p]}| > 1. \end{aligned}$$

Note that foci are points with locally stationary distance to a subset of L .

If $G = \{id\}$ and $M = \mathbb{R}^m$, then these definitions agree with Bhattacharya and Patrangenaru ((2003), p.2 and p.12). Also for $M = \mathbb{R}^m$, foci have been introduced as ‘‘focal points’’ by Milnor ((1969), p.32); Hastie and Stuetzle ((1989), p.514) call the focal points of a one-dimensional submanifold of Euclidean space the *ambiguity set*. We first give some illustrations in the Euclidean plane.

1. The center of a circle is both a focus and a focal point to that circle; the ‘‘foci’’ of a non-circular ellipse are its foci, the open line segment between them constitute its focal points, none of which is a focus.

2. Consider a suitably smoothed version of the union of a circular segment and a line segment

$$\left\{ -\frac{1}{\sqrt{2}} \right\} \times \left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right] \cup \left\{ (x, y) \in \mathbb{R}^2 : x \geq -\frac{1}{\sqrt{2}}, x^2 + y^2 = 1 \right\}.$$

Then the origin is its focus, the minimal distance to the origin, however, is attained at $(-1/\sqrt{2}, 0)$, from which the origin is reached by the endpoint map with non-singular derivative.

3. In order to see that the set of foci of a closed manifold is not necessarily closed, consider a suitably smoothed version M of the union $\cup_{n=2}^{\infty} (K_n \cup L_n)$ of circular segments K_n and line segments L_n defined by

$$\begin{aligned} K_n & := \left\{ (x, y) \in \mathbb{R}^2 : \left(x - \frac{1}{n}\right)^2 + y^2 = R_n^2, x > 0, |y| \leq 2 \right\}, \quad R_n := n - \frac{1}{n} \\ L_n & := \begin{cases} \left\{ (x, -2) : \frac{1}{n} + \sqrt{R_n^2 - 4} \leq x \leq \frac{1}{n+1} + \sqrt{R_{n+1}^2 - 4} \right\} & n \text{ odd} \\ \left\{ (x, 2) : \frac{1}{n} + \sqrt{R_n^2 - 4} \leq x \leq \frac{1}{n+1} + \sqrt{R_{n+1}^2 - 4} \right\} & n \text{ even.} \end{cases} \end{aligned}$$

Then, all $p_n = (1/n, 0)$ with $n \geq 2$ are foci of M ; the origin, their limit point, however is not a focus.

We are concerned with *minimizing foci* which form a closed subset for closed manifolds as we shall see.

Definition A.4. *A focus p of L will be called minimizing if there is $(p', v) \in NL$ such that $\exp_{p'}(v) = p$, $(\mathrm{d}\phi)_{(p', v)}$ is singular and $d_Q(p, L) = \|v\|$.*

The following is a generalization of Theorem 3.2 of Bhattacharya and Patrangenaru ((2003), p.12), which is a generalization of Proposition 6 of Hastie and Stuetzle ((1989), p.515).

Theorem A.5. *Let G be a compact Lie group acting isometrically on a complete finite-dimensional Riemannian manifold M , and let L be an arbitrary submanifold of M , K an arbitrary generalized submanifold of $Q = M/G$. Then*

- (a) *the set of foci and focal points of L has measure zero in M ,*
- (b) *if L is closed then the set of minimizing foci and focal points of L is closed in M ,*
- (c) *the set of focal points of K has measure zero in Q .*

Proof. Throughout the proof we use the notation introduced above. Moreover let

- \mathcal{L}^0 be the set of foci of L ,
- \mathcal{L}^{00} be the set of minimizing foci of L ,
- \mathcal{L} be set the of focal points of L , and
- \mathcal{K} be set the of focal points of K .

From the following Claims I and II we obtain the assertion (a). Claims III and IV yield assertion (b), and Claim V assertion (c).

Claim I: \mathcal{L}^0 is of measure zero. By definition, \mathcal{L}^0 is the set of critical points of the endpoint map $\phi : NL \rightarrow M$. As in Milnor ((1969), p.33), Sard's theorem ensures that the critical points have measure zero in M .

Claim II: $\mathcal{L}^1 := \mathcal{L} \setminus \mathcal{L}^0$ is of measure zero. From every focus $p \in \mathcal{L}$ we have the set of orthogonal projections L_p . Conversely, for every point $p' \in L_p \subset M$, there is a unique normal vector $v = v(p, p')$ such that $p = \exp_{p'}(v)$. If $p' \in L_{p_1} \cap L_{p_2}$ for focal points $p_1 \neq p_2$, then obviously $v(p_1, p')$ and $v(p_2, p')$ cannot be collinear. Hence, there is a subset $A \subset CL$ of the cylindrical manifold

$$CL := \{(q, v) \in NL : \|v\| = 1\}$$

around L in NL and a mapping $\psi : A \rightarrow NL, (q, v) \rightarrow (q, t(q, v)v)$ such that $\chi := \phi \circ \psi : A \rightarrow \mathcal{L}^1$ is surjective. We show that χ is locally homeomorphic. Then, $\mathcal{L}^1 \subset M$ is the locally homeomorphic image of a subset A of a set of measure zero CL in NL with $\dim(NL) = \dim(M)$; hence \mathcal{L}^1 is a set of measure zero in M .

Continuity of χ follows from continuity of t : let $A \ni (q_n, v_n) \rightarrow (q, v) \in A$ and $t_n = t_n(q_n, v_n) \rightarrow t_0$ with focal points $\mathcal{L}^1 \ni p_n := \exp_{q_n}(t_n v_n) \rightarrow \exp_q(t_0 v) = p$. By continuity of distance, $d_M(p, q) = d_M(p, L)$. By hypothesis $\tilde{p} = \exp_q(t(q, v)v)$ is focal with $d_M(\tilde{p}, L) = d(\tilde{p}, q)$. This yields $t_0 = t(q, v)$.

Moreover, $\chi^{-1} = \rho \circ \phi^{-1}|_{\mathcal{L}^1}$ is locally well defined and continuous as it is the composition of the continuous projection

$$\rho : \{(q, v) \in NL : \|v\| \neq 0\} \rightarrow CL, \quad (q, v) \mapsto \left(q, \frac{v}{\|v\|} \right),$$

and a locally diffeomorphic inverse branch ϕ^{-1} around \mathcal{L}^1 ; by hypothesis we stay away from the singularity set \mathcal{L}^0 of ϕ .

Claim III: \mathcal{L}^{00} is closed. Consider sequences $p_n \in \mathcal{L}^{00}$, $p'_n \in L$, and $v_n \in T_{p'_n}L$ such that $p_n = \exp_{p'_n}(v_n)$ and $d\phi_{(p'_n, v_n)}$ is singular. If $p_n \rightarrow p \in M$ then, as we are considering minimizing foci only, $\|v_n\|$ is bounded. As a consequence, p'_n has a point of accumulation $p' \in M$ which is in L if the latter is closed. By continuity there is a $v \in N_{p'}L$ with $\exp_{p'}(v) = p$ and $\|v\| = d_M(p, L)$. Again by continuity, $d\phi_{q, v}$ is singular, proving that $p \in \mathcal{L}^{00}$.

Claim IV: $\mathcal{L}^{00} \cup \mathcal{L}$ is closed. In order to complete the argument that the set of minimizing foci and focal points is closed, we consider a non-focal limit point $p \in M$ of a sequence of focal points $p_n \in M$ and show that it is a minimizing focus. Indeed if $q_n, q'_n \in L_{p_n}$, $q_n \neq q'_n$, we may assume that $q_n, q'_n \in L$ have a common point of accumulation $q \in L$ (since p is non-focal) and that

$$\exp_{q_n}(t_n v_n) = p_n = \exp_{q'_n}(t_n v'_n) \quad (21)$$

for suitable unit length $v_n \in N_{q_n}L$, $v'_n \in N_{q'_n}L$, and $t_n = d_M(q_n, L) > 0$. Again, t_n has a point of accumulation t and v_n, v'_n have a point of accumulation v such that

$$\exp_q(tv) = p, t = d_M(p, L).$$

Hence, from (21) we infer at once that $d\phi_{(p, tv)}$ is singular and thus p is a minimizing focus.

Claim V: \mathcal{K} is of measure zero in Q . Since there are only countably many orbit types in M , by hypothesis there is an index set $I \subset \mathbb{N}$ such that $K = \cup_{i \in I} K_i$, with submanifolds $K_i := K \cap Q^{(H_i)}$ of $Q^{(H_i)}$ and closed subgroups H_i of G for $i \in I$. As noted in the beginning of this section, the inverse projection $L_i := \pi^{-1}(K_i)$ of each of these is a submanifold of M . Moreover, every point $p \in [p]$, where $[p]$ is focal for K , is focal for $L_i \cup L_j$ with suitable indices $i, j \in I$. As the argument of claim II is local in nature, it can be applied to every L_i with

$$\tilde{\mathcal{L}}_i^1 := \{p \in \mathcal{L}^1 \setminus \mathcal{L}_i^0 : L_p \cap L_i \neq \emptyset\}$$

instead of $\mathcal{L}_i^1 \setminus \mathcal{L}_i^0$; here $L = \cup_{i \in I} L^i$ and $\mathcal{L}_i^0, \mathcal{L}_i^1$ denote the foci and focal points, resp. of L_i . This yields that the set of focal points for $L = \cup_{i \in I} L_i$ is of measure zero in M . As this is the inverse projection of \mathcal{K} , the latter is of measure zero in Q . \square

Appendix B: Tree-Stem Data Set

Landmarks	1	2	3	4
Tree 1	27	14	14	14
	0	0	0.05662873	-0.09531028
	0	0	-0.0674875	0.03469011
Tree 2	29.7	17.6	17.6	17.6
	0	0	-0.07479217	0.0587687
	0	0	-0.04318128	0.1017904
Tree 3	32.5	18.2	18.2	18.2
	0	0	0.1427532	-0.1581646
	0	0	$-3.496329e - 17$	0.09131639
Tree 4	28	18	18	18
	0	0	0.05264516	0.01673678
	0	0	-0.06274006	0.094919
Tree 5	24.3	16.1	16.1	16.1
	0	0	-0.0591466	0.04722901
	0	0	-0.03414831	0.08180305

Table 5: *Tetrahedral configurations from the stems of 5 Douglas fir trees. Units are given in meters.*

Acknowledgment

We are indebted to the editors and to the three anonymous referees whose encouraging comments helped to improve this paper considerably. We also thank our colleagues from the Institute for Forest Biometry and Informatics at the University of Göttingen. The first author would like to express his gratitude for many fruitful discussions with Werner Ballmann, Andreas Lytchak, Walter Oevel, Anita Schöbel and Victor Vuletescu. He also gratefully acknowledges support by DFG grant MU 1230/10-1 and by Volkswagen Stiftung. The second author gratefully acknowledges support by DFG Research Training Group 1023 and by the German Federal Ministry of Education and Research, Grant 03MUPAH6.

Bibliography

- Abraham, R., and Marsden, J. E. (1978). *Foundations of Mechanics*, 2nd Edition. Benjamin-Cummings.
- Ambartzumian, R. V. (1990). *Factorization, Calculus and Geometric Probability*. Cambridge University Press.
- Bhattacharya, R. N., and Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds I. *Ann. Statist.* **31**, 1–29.
- Bhattacharya, R. N., and Patrangenaru, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds II. *Ann. Statist.* **33**, 1225–1259.
- Blum, H., and Nagel, R. N. (1978). Shape description using weighted symmetric axis features. *Pattern Recognition* **10**, 167–180.
- Bookstein, F. L. (1978). *The Measurement of Biological Shape and Shape Change*, 2nd Edition. Vol. 24 of *Lecture Notes in Biomathematics*. Springer-Verlag, New York.
- Borzellino, J., Jordan-Squire, C., Petrics, G., and Sullivan, D. (2007). On the existence of infinitely many closed geodesics on orbifolds of revolution. Submitted.
- Bredon, G. E. (1972). *Introduction to Compact Transformation Groups*. Vol. 46 of *Pure and Applied Mathematics*. Academic Press.
- Bubenik, P., and Kim, P. T. (2007). A statistical approach to persistent homology. *Homology, Homotopy and Applications* **9**, 337–362.
- Cootes, T. F., Taylor, C. J., H., C. D., and Graham, J. (1992). Training models of shape from sets of examples. In: *British Machine Vision Conference*, pp. 9–18, Springer, Berlin.
- de Silva, V., and Carlsson, G. (2004). Topological estimation using witness complexes. In: *SPBG04 Symposium on Point-Based Graphics*, pp. 157–166.
- Dryden, I. L., and Mardia, K. V. (1998). *Statistical Shape Analysis*. Wiley, Chichester.

- Fletcher, P. T., and Joshi, S. C. (2007). Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing* **87**, 250–262.
- Fletcher, P. T., Lu, C., Pizer, S. M., and Joshi, S. C. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging* **23**, 995–1005.
- Fuchs, M., and Scherzer, O. (2007). Regularized reconstruction of shapes with statistical a priori knowledge. Technical Report 51, FSP 092.
- Gaffrey, D., and Sloboda, B. (2000). Dynamik der Stammorphologie: Ansätze zur Quantifizierung und Modellierung der Assimilatakollation. In: *Proceedings Deutscher Verband Forstl. Forschungsanstalten, Sektion Forstl. Biometrie und Informatik. 12. Tagung*, pp. 7–26, Göttingen.
- Goodall, C. R. (1991). Procrustes methods in the statistical analysis of shape (with discussion). *Journal of the Royal Statistical Society B* **53**, 285–339.
- Goodall, C. R., and Mardia, K. V. (1999). Projective shape analysis. *J. Graphical and Computational Statist.* **8**, 143–168.
- Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika* **40**, 33–51.
- Hastie, T., and Stuetzle, W. (1989). Principal curves. *Journal of the American Statistical Association* **84**, 502–516.
- Helgason, S. (1962). *Differential Geometry and Symmetric Spaces*. Academic Press, New York.
- Hendriks, H., and Landsman, Z. (1998). Mean location and sample mean location on manifolds: asymptotics, tests, confidence regions. *Journal of Multivariate Analysis* **67**, 227–243.
- Hendriks, H., Landsman, Z., and Ruymgaart, F. (1996). Asymptotic behaviour of sample mean direction for spheres. *Journal of Multivariate Analysis* **59**, 141–152.
- Hodson, F. R., Sneath, P. H., and Doran, J. E. (1966). Some experiments in the numerical analysis of archeological data. *Biometrika* **53**, 411–324.
- Hotz, T., Huckemann, S., Gaffrey, D., Munk, A., and Sloboda, B. (2007). Keeping intrinsic PCA simple: log-linear shape spaces with an application to the temporal evolution of tree stems. *Journal of the Royal Statistical Society C*. In revision.
- Hsieh, W. W. (2007). Nonlinear principal component analysis of noisy data. *Neural Netw.* **20**, 434–443.
- Huckemann, S., and Hotz, T. (2007). Principal component geodesics for planar shape spaces. *Journal of Multivariate Analysis*. In revision.

- Huckemann, S., and Ziezold, H. (2006). Principal component analysis for Riemannian manifolds with an application to triangular shape spaces. *Adv. Appl. Prob.* **38**, 299–319.
- Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics* **XXX**, 509–541.
- Kendall, D. G. (1984). Shape manifolds, procrustean metrics and complex projective spaces. *Bull. Lond. Math. Soc.* **16**, 81–121.
- Kendall, D. G., Barden, D., Carne, T. K., and Le, H. (1999). *Shape and Shape Theory*. Wiley, Chichester.
- Kent, J. (1994). The complex Bingham distribution and shape analysis. *Journal of the Royal Statistical Society B* **56**, 285–299.
- Kim, P. T., and Koo, J.-Y. (2005). Statistical inverse problems on manifolds. *Journal of Fourier Analysis and Applications* **11**, 639–653.
- Klassen, E., Srivastava, A., Mio, W., and Joshi, S. (2004). Analysis on planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, 372–383.
- Kneip, A., and Utikal, K. J. (2001). Inference for density families using functional principal component analysis (with discussion). *J. Am. Stat. Assoc.* **96**, 519–542.
- Kobayashi, S., and Nomizu, K. (1969). *Foundations of Differential Geometry, Vol. II*. Wiley, Chichester.
- Krim, H., and Yezzi, A. J. J. E. (2006). *Statistics and Analysis of Shapes. Modeling and Simulation in Science, Engineering and Technology*. Birkhäuser, Boston.
- Lang, S. (1999). *Fundamentals of Differential Geometry*. Springer.
- Le, H. (2001). Locating Fréchet means with an application to shape spaces. *Adv. Appl. Prob.* **33**, 324–338.
- Le, H., and Kume, A. (2000). Detection of shape changes in biological features. *Journal of Microscopy* **200**, 140–147.
- Mardia, K., and Patrangenaru, V. (2001). On affine and projective shape data analysis. In: *Functional and Spatial Data Analysis, Proceedings of the 20th LASR Workshop* (Eds: K.V. Mardia and R.G. Aykroyd), 39–45.
- Michor, P. W., and Mumford, D. (2006). Riemannian geometries on spaces of plane curves. *J. of the European Math. Soc.* **8**, 1–48.

- Miller, M. I., Trounev, A., and Younes, L. (2006). Geodesic shooting for computational anatomy. *J. Math. Imaging Vis.* **24**, 209–228.
- Milnor, J. W. (1969). *Morse Theory*. Princeton University Press, Princeton, 3rd printing with corrections.
- Munk, A., Paige, R., Pang, J., Patrangenaru, V., and Ruymgaart, F. (2007). The one- and multi-sample problem for functional data with application to projective shape analysis. *Journal of Multivariate Analysis*, 815–833.
- Palais, R. S. (1960). Slices and equivariant embeddings. In: *Seminar on Transformation Groups*, Borel, A. (Ed.), no. 46 in *Annals of Mathematics Studies*, pp. 101–115, Princeton Univ. Press, Princeton NJ.
- Palais, R. S. (1961). On the existence of slices for actions of non-compact Lie groups. *Ann. Math.* 2nd Ser. **73**, 295–323.
- Pizer, S. M., Siddiqi, K., Székely, G., Damon, J. N., and Zucker, S. W. (2003). Multiscale medial loci and their properties. *Int. J. Comput. Vision* **55**, 155–179.
- Schmidt, F. R., Clausen, M., and Cremers, D. (2006). Shape matching by variational computation of geodesics on a manifold. In: *Pattern Recognition (Proc. DAGM)*, Vol. 4174 of LNCS, pp. 142–151, Springer, Berlin, Germany.
- Schmidt, F. R., Töppe, E., Cremers, D., and Boykov, Y. (2007). Intrinsic mean for semi-metrical shape retrieval via graph cuts. In: *DAGM-Symposium*, Hamprecht, F. A., Schnörr, C., and Jähne, B. (Eds.), Vol. 4713 of *Lecture Notes in Computer Science*, pp. 446–455, Springer.
- Small, C. G. (1996). *The Statistical Theory of Shape*. Springer-Verlag, New York.
- Ziezold, H. (1977). Expected figures and a strong law of large numbers for random elements in quasi-metric spaces. *Trans. 7th Prague Conf. Inf. Theory, Stat. Dec. Func., Random Processes A*, 591–602.
- Zwald, L., Bousquet, O., and Blanchard, G. (2004). Learning theory. In: *COLT*. Vol 3120 of *Lecture Notes in Computer Science*, Shawe-Taylor, J., and Singer, Y. (Eds.), pp. 594–608, Springer.
- Stephan Huckemann: Institute for Mathematical Stochastics, Georgia Augusta Universität Göttingen, Maschmühlenweg 8-10, D-37073 Göttingen,
E-mail: huckeman@math.uni-goettingen.de
- Thomas Hotz: Institute for Mathematical Stochastics, Georgia Augusta Universität Göttingen,
E-mail: hotz@math.uni-goettingen.de
- Axel Munk: Institute for Mathematical Stochastics, Georgia Augusta Universität Göttingen,

E-mail: munk@math.uni-goettingen.de

Statistica Sinica