Ann Inst Stat Math (2012) 64:1227–1259
DOI 10.1007/s10463-012-0352-2

On the meaning of mean shape: manifold stability, locus and the two sample test

Stephan F. Huckemann

Received: 27 May 2011 / Revised: 26 September 2011 / Published online: 13 March 2012

Abstract Various concepts of mean shape previously unrelated in the literature are brought into relation. In particular, for non-manifolds, such as Kendall's 3D shape space, this paper answers the question, for which means one may apply a two-sample test. The answer is positive if intrinsic or Ziezold means are used. The underlying general result of manifold stability of a mean on a shape space, the quotient due to an proper and isometric action of a Lie group on a Riemannian manifold, blends the slice theorem from differential geometry with the statistics of shape. For 3D Procrustes means, however, a counterexample is given. To further elucidate on subtleties of means, for spheres and Kendall's shape spaces, a first-order relationship between intrinsic, residual/Procrustean and extrinsic/Ziezold means is derived stating that for high concentration the latter approximately divides the (generalized) geodesic segment between the former two by the ratio 1:3. This fact, consequences of coordinate choices for the power of tests and other details, e.g. that extrinsic Schoenberg means may increase dimension are discussed and illustrated by simulations and exemplary datasets.

Keywords Intrinsic mean · Extrinsic mean · Procrustes mean · Schoenberg mean · Ziezold mean · Shape spaces · Proper Lie group action · Slice theorem · Horizontal lift · Stratified spaces

1 Introduction

The analysis of shape may be counted among the very early activities of mankind; be it for representation on cultural artefacts, or for morphological, biological and medical

S. F. Huckemann (✉)
Institute for Mathematical Stochastics, University of Göttingen, Goldschmidstr. 5–7, Göttingen, Germany
e-mail: huckeman@math.uni-goettingen.de

Table 1 Three fundamental types of means on a shape space (right column) and their horizontal lifts to the respective manifold (left column)

Manifold means	Shape means
Intrinsic	Intrinsic
Extrinsic	Ziezold
residual	Procrustean

applications. In modern days shape analysis is gaining increased momentum in computer vision, image analysis, biomedicine and many other fields. For a recent overview (cf. Krim and Yezzi 2006).

A *shape space* can be viewed as the quotient of a Riemannian manifold—e.g. the pre-shape sphere of centered unit size landmark configurations—modulo the isometric and proper action of a Lie group (cf. Bredon 1972), conveying shape equivalence—e.g. the group of rotations (cf. Kendall et al. 1999, Chapter 11). Thus, it carries the canonical quotient structure of a union of manifold strata of different dimensions, which give in general a Riemannian *manifold part*—possibly with singularities comprising the *non-manifold part* of *non-regular shapes* at some of which sectional curvatures may tend to infinity (cf. Kendall et al. 1999, Chapter 7.3) as well as Huckemann et al. (2010b).

In a Euclidean space, there is a clear and unique concept of a mean in terms of least squares minimization: the arithmetic average. Generalizing to manifolds, however, the concept of expectation, average or *mean* is surprisingly non trivial and not at all canonical. In fact, it resulted in an overwhelming number of different concepts of means, each defined by a specific concept of a distance, all of which are identical for the Euclidean distance in a Euclidean space. More precisely, with every embedding in a Euclidean space come specific *extrinsic* and *residual means* and with every Riemannian structure comes a specific *intrinsic mean*. Furthermore, due to the non-Euclidean geometry, local minimizers introduced as *Karcher means* by Kendall (1990) may be different from global minimizers called *Fréchet means* by Ziezold (1977), and, neither ones are necessarily unique. Nonetheless, carrying statistics over to manifolds, strong consistency by Ziezold (1977), Bhattacharya and Patrangenaru (2003) and under suitable conditions, central limit theorems (CLTs) for such means have been derived by Jupp (1988), Hendriks and Landsman (1996, 1998), Bhattacharya and Patrangenaru (2005) as well as Huckemann (2011). On shape spaces, various other concepts of means have been introduced, e.g. the famous *Procrustes means* (cf. Gower 1975; Ziezold 1977; Dryden and Mardia 1998). As we show here, these means are related to the above ones via a *horizontal lifting* from the bottom quotient to the top manifold (cf. Table 1. In particular, since there are many—and often confusing—variants of Procrustes means in the literature this paper takes the advantage of a nonstandard geometric viewpoint and introduces the terminology of *Procrustean means* standing for inheritance from residual means.

The CLTs quoted above assume a manifold structure locally relating to a Euclidean space; namely, asserting that under scaling by the square root of sample size, in a local chart sample means are asymptotically normally distributed. As in the Euclidean case this allows for *one-sample* and *two-sample* tests. This argument fails, however, if sample means converge to a singularity. For such cases the asymptotic distribution is not known in general, in consequence there are no one- and two-sample tests available. On tree spaces (e.g. Billera et al. 2001), ongoing work has come up with a "sticky CLT on the spine of the open book" (Hotz et al. 2011). In particular, intrinsic means on such spaces tend to lie on the singular part. For shape spaces, this leads to the question under which conditions it can be guaranteed that a mean shape sticks not to the singular part but lies "stably" on the manifold part.

Due to strong consistency, for a one-sample test for a specific mean shape on the manifold part, it may be assumed that sample means eventually lie on the manifold part as well, thus making the above cited CLTs available. A two- and a multi-sample test for common mean shape, however, could not be justified to date because of a lacking result on the following manifold stability.

Definition 1 A mean shape enjoys *manifold stability* if it is assumed on the manifold part for any random shape assuming the manifold part with non-zero probability.

A key result of this paper establishes manifold stability for intrinsic and Ziezold means under the following condition.

Condition 1 *On the non-manifold part the distribution of the random shape contains at most countably many point masses.*

Since the non-manifold part is a null-set (e.g. Bredon 1972) under the projection of the Riemannian volume, this condition covers most realistic cases.

We develop the corresponding theory for a general shape space quotient based on lifting a distribution on the shape space to the pre-shape space and subsequently exploiting the fact that intrinsic means are zeroes of an integral involving the Riemann exponential. The similar argument can be applied to Ziezold means, but not to Procrustean means. More specifically, we develop the notion of a *measurable horizontal lift* of the shape space except for its *quotient cut locus* (introduced as well) to the pre-shape space. This requires the geometric concept of *tubular neighborhoods admitting slices*.

Curiously, the result applied to the finite dimensional subspaces exhausting the quotient shape space of closed planar curves with arbitrary initial point introduced by Zahn and Roskies (1972) and further studied by Klassen et al. (2004), gives that the shape of the circle, since it is a singularity, can never be an intrinsic shape mean of non-circular curves.

As a second curiosity, 3D full Procrustes means do not enjoy manifold stability in general, a counterexample involving low concentration is given. This is due to the fact that for low concentration, full Procrustes means may be 'blinder' in comparison to intrinsic and Ziezold means to distributional changes far away from a mode. Included in this context is also a discussion of the *Schoenberg means*, recently introduced by Bandulasiri and Patrangenaru (2005) as well as by Dryden et al. (2008) for the non-manifold Kendall reflection shape spaces, which in the ambient space, also allow for a CLT. Schoenberg means, as demonstrated, however, may feature 'blindness' in comparison to intrinsic and Ziezold means, with respect to changes in the distribution of nearly degenerate shapes. In a simulation we show that these features render Schoenberg means less effective for a discrimination involving degenerate or nearly degenerate shapes.

As a third curiosity, for spheres and Kendall's shape spaces, it is shown that, given uniqueness, with order of concentration, the (generalized) geodesic segment between the intrinsic mean and the residual/Procrustean mean is in approximation divided by the extrinsic/Ziezold mean by the ratio 1:3. This first order relationship can be readily observed in existing data sets. In particular, this result supports the conjecture that Procrustean means of sufficiently concentrated distributions enjoy stability as well.

This paper is structured as follows. For the convenience of the reader the following Sect. 2 is a self- contained account on manifold stability for Procrustes and other means on Kendall's shape spaces which can be read alone. This section is followed by a general classification of concepts of means on general shape spaces in Sect. 3. The rather technical Sect. 4 develops horizontal lifting and establishes manifold stability, technical proofs are deferred to the Appendix. In Sect. 5, extrinsic Schoenberg means are discussed and Sect. 6 tackles local effects of curvature on spheres and Kendall's shape spaces. Section 7 illustrates practical consequences using classical data-sets as well as simulations. Note that lacking stability does not affect the validity of the Strong Law, on which the considerations on asymptotic distance in Sects. 6 and 7 are based.

An R-package for all of the computations performed is provided online: Huckemann (2010).

2 Stability of means on Kendall's shape spaces

In the statistical analysis of similarity shapes based on landmark configurations, geometrical m-dimensional objects (usually $m = 2, 3$) are studied by placing $k > m$ *landmarks* at specific locations of each object. Each object is then described by a matrix in the space $M(m, k)$ of $m \times k$ matrices, each of the k columns denoting an m-dimensional landmark vector. $\langle x, y\rangle := \operatorname{tr}(xy^T)$ denotes the usual inner product with norm $\|x\| = \sqrt{\langle x, x\rangle}$. For convenience and without loss of generality for the considerations below, only *centered* configurations are considered. Centering can be achieved by multiplying with a sub-Helmert matrix $\mathcal{H} \in M(k, k-1)$ from the right, yielding a configuration $x\mathcal{H}$ in $M(m, k-1)$. For this and other centering methods (cf. Dryden and Mardia 1998, Chapter 2). Excluding also all configurations with all landmarks coinciding gives the space of *configurations*

$$F_m^k := M(m, k-1) \setminus \{0\}.$$

Since only the similarity shape is of concern, we may assume that all configurations are contained in the unit sphere $S_m^k := \{x \in F_m^k : \|x\| = 1\}$ called the *pre-shape sphere*. With $O(m) := \{g \in M(m, m) : g^T g = e\}$ denoting the orthogonal group, $e := \operatorname{diag}(1, \dots, 1)$ the unit matrix, $\widetilde{e} := \operatorname{diag}(-1, 1, \dots, 1)$ and $SO(m) := \{g \in O(m) : \det(g) = 1\}$ the special orthogonal group, *Kendall's shape space* is then the canonical quotient

$$\Sigma_m^k := S_m^k / SO(m) = \{[x] : x \in S_m^k\} \text{ with the } \textit{orbit } [x] = \{gx : g \in SO(m)\}.$$

In some applications reflections are also filtered out giving *Kendall's reflection shape space*

$$R\Sigma_m^k := \Sigma_m^k/\{e, \widetilde{e}\} = S_m^k/O(m).$$

For $1 \le j < m < k$ consider the isometric embedding

$$S_j^k \hookrightarrow S_m^k : x \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix} \tag{1}$$

giving rise to a canonical embedding $R\Sigma_j^k \hookrightarrow \Sigma_m^k$ which is isometric with respect to

$$\begin{array}{lll} \textit{the intrinsic distance} & \rho^{(i)}(x, x') := \min\limits_{g \in G} & \arccos\langle gx, x'\rangle\,, \\ \textit{the Ziezold distance} & \rho^{(z)}(x, x') := \min\limits_{g \in G} & \sqrt{2 - 2\langle gx, x'\rangle} \\ \textit{and the Procrustean distance} & \rho^{(p)}(x, x') := \min\limits_{\substack{g \in G \\ \langle gx, x'\rangle \ge 0}} & \sqrt{1 - \langle gx, x'\rangle^2} \end{array}$$

with $G = SO(m)$ for Σ_m^k and $G = O(j)$ for $R\Sigma_j^k$ (cf. Sect. 3) (Kendall et al. 1999, p. 29) and also Remark 1 below.

We say that a configuration in $\mathbb{R}^m$ is *j-dimensional*, or more precisely *non-degenerate j-dimensional* if its preshape $x \in S_m^k$ is of rank j. Moreover, for $j \ge 3$ the shape spaces Σ_j^k and $R\Sigma_j^k$ decompose into a *manifold part* (cf. Sects. 3, 5) of *regular shapes*

$$(\Sigma_j^k)^* = \{[x] \in \Sigma_j^k : \operatorname{rank}(x) \ge j - 1\} \text{ and } (R\Sigma_j^k)^* = \{[x] \in R\Sigma_j^k : \operatorname{rank}(x) = j\}\,,$$

respectively, given by the shapes corresponding to configurations of at least dimension $j-1$ and j, respectively and a non void part of singular shapes corresponding to lower dimensional configurations, respectively.

Given a random shape $[X] \in \Sigma = \Sigma_m^k$ or $R\Sigma_j^k$, various concepts of *expected shapes* are possible. The sets of minimizers of the following expectations are called

$$\begin{array}{ll} \textit{intrinsic means:} & \operatorname*{argmin}\limits_{q \in \Sigma} \mathbb{E}\big(\rho^{(p)}(q, [X])^2\big)\,, \\ \textit{Ziezold means:} & \operatorname*{argmin}\limits_{q \in \Sigma} \mathbb{E}\big(\rho^{(z)}(q, [X])^2\big)\,, \text{ and} \\ \textit{full Procrustes means:} & \operatorname*{argmin}\limits_{q \in \Sigma} \mathbb{E}\big(\rho^{(p)}(q, [X])^2\big). \end{array}$$

We say that a mean is *unique* if the corresponding set contains one element only. A detailed discussion of these and more concept of means can be found in Sect. 3.

The proofs of the following two Theorems can be found in the Appendix.

Theorem 1 *Suppose that $[X]$ is a random shape on Σ_m^k assuming shapes in $R\Sigma_j^k$ ($1 \le j < m < k$) with probability one. Then every full Procrustes mean shape of $[X]$ and every unique intrinsic or Ziezold mean shape assuming the non-manifold part $(R\Sigma_j^k) \setminus (R\Sigma_j^k)^*$ only with at most countably many point masses corresponds to a configuration of dimension less than or equal to j.*

The following theorem is the application of the key result applied to Kendall's shape spaces.

Theorem 2 (Stability theorem for intrinsic and Ziezold means) *Let $[X]$ be a random shape on Σ_m^k, $0 < m < k$, with unique intrinsic or Ziezold mean shape $[\mu] \in \Sigma_m^k$, $[\mu] \in S_m^k$ and let $1 \le j \le m$ be the maximal dimension of configurations of shapes assumed by X with non-zero probability. Suppose moreover that shapes of configurations of strictly lower dimensions are assumed with at most countably many point masses.*

1. *If $j < m$ then $[\mu]$ corresponds to a non-degenerate j-dimensional configuration.*
2. *If $j = m$ then $[\mu]$ corresponds to a non-degenerate configuration of dimension $m - 1$ or m.*

Remark 1 The result of Theorem 2 is sharp. To see this, consider for $\alpha > \beta > 0$, $\alpha^2 + \beta^2 = 1$ the pre-shapes

$$x = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \end{pmatrix}, \quad y = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & -\beta & 0 \end{pmatrix} \text{ and } z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in S_2^4.$$

Then x and y correspond to non-degenerate two-dimensional quadrilateral configurations while z corresponds to a one-dimensional (collinear) quadrilateral. Still, $[z]$ is regular in Σ_2^4 and it is the intrinsic and Ziezold mean of $[x]$ and $[y]$ in Σ_2^4. Under the embedding $R\Sigma_2^4 \hookrightarrow \Sigma_3^4$ we have the pre-shapes

$$x' = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad y' = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & -\beta & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } z' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in S_3^4.$$

Just as $[x] = [y]$ in $R\Sigma_2^4$ so do x' and y' have regular and identical shape in Σ_3^4. However, $[z']$ is not regular and it is not the intrinsic or Ziezold mean in Σ_3^4.

3 Fundamental types of means

In the previous section we introduced Kendall's *shape* and *reflection shape space* based on invariance under similarity transformations and, including reflections, respectively. Invariance under congruence transformations only leads to Kendall's *size-and-shape space*. More generally in image analysis, invariance may also be considered under the affine or projective group (cf. Mardia and Patrangenaru 2001, 2005). A different yet also very popular set of shape spaces for two-dimensional configurations modulo the group of similarities has been introduced by Zahn and Roskies (1972). Instead of building on a finite dimensional Euclidean matrix space modeling landmarks, the basic ingredient of these spaces modeling closed planar unit speed curves is the infinite dimensional Hilbert space of Fourier series (cf. Klassen et al. 2004). In practice for numerical computations, only finitely many Fourier coefficients are considered.

To start with, a shape space is a metric space (Q, d). For this entire paper suppose that $X, X_1, X_2, \ldots$ are i.i.d. random elements mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to (Q, d) equipped with its self understood Borel σ-field. Here and in the following, *measurable* will refer to the corresponding Borel σ-algebras, respectively. Moreover, denote by $\mathbb{E}(Y)$ the classical expected value of a random vector Y on a D-dimensional Euclidean space $\mathbb{R}^D$, if existent.

Definition 2 For a continuous function $\rho : Q \times Q \to [0, \infty)$ define the *set of population Fréchet ρ-means* by

$$E^{(\rho)}(X) = \operatorname*{argmin}_{\mu \in Q} \mathbb{E}\big(\rho(X, \mu)^2\big).$$

For $\omega \in \Omega$ denote the *set of sample Fréchet ρ-means* by

$$E_n^{(\rho)}(\omega) = \operatorname*{argmin}_{\mu \in Q} \sum_{j=1}^{n} \rho\big(X_j(\omega), \mu\big)^2.$$

By continuity of ρ, the ρ-means are closed sets, additionally, sample ρ-means are random sets, all of which may be empty. For our purpose here, we rely on the definition of *random closed sets* as introduced and studied by Choquet (1954), Kendall (1974) and Matheron (1975). Since their original definition for $\rho = d$ by Fréchet (1948) such means have found much interest.

Intrinsic means. Independently, for a connected Riemannian manifold with geodesic distance $\rho^{(i)}$, Kobayashi and Nomizu (1969) defined the corresponding means as *centers of gravity*. They are nowadays also well known as *intrinsic means* by Bhattacharya and Patrangenaru (2003, 2005).

Extrinsic means. With respect to the chordal or *extrinsic metric* $\rho^{(e)}$ due to an embedding of a Riemannian manifold in an ambient Euclidean space, Fréchet ρ-means have been called *mean locations* by Hendriks and Landsman (1996) or *extrinsic means* by Bhattacharya and Patrangenaru (2003).

More precisely, let $Q = M \subset \mathbb{R}^D$ be a complete Riemannian manifold embedded in a Euclidean space $\mathbb{R}^D$ with standard inner product $\langle x, y\rangle$, $\|x\| = \sqrt{\langle x, x\rangle}$, $\rho^{(e)}(x, y) = \|x - y\|$ and let $\Phi : \mathbb{R}^D \to M$ denote the orthogonal projection, $\Phi(x) = \operatorname{argmin}_{p \in M} \|x - p\|$. For any Riemannian manifold an embedding that is even isometric can be found for D sufficiently large, see Nash (1956). Due to an extension of Sard's Theorem by (Bhattacharya and Patrangenaru 2003, p.12) for a closed manifold, Φ is univalent up to a set of Lebesgue measure zero. Then the set of extrinsic means is given by the set of images $\Phi\big(\mathbb{E}(Y)\big)$ where Y denotes X viewed as taking values in $\mathbb{R}^D$ (cf. Bhattacharya and Patrangenaru 2003).

Residual means. In this context, setting $\rho^{(r)}(p, p') = \|d\Phi_{p'}(p - p')\|$ $(p, p' \in M)$ with the derivative $d\Phi_{p'}$ at p' yielding the orthogonal projection to the embedded tangent space $T_{p'}\mathbb{R}^D \to T_{p'}M \subset T_{p'}\mathbb{R}^D$, call the corresponding mean sets $E^{(\rho^{(r)})}(X)$ and $E_n^{(\rho^{(r)})}(\omega)$, the sets of *residual population means* and *residual sample means*, respectively. For two-spheres, $\rho^{(r)}(p, p')$ has been studied under the name of *crude residuals* by Jupp (1988). On unit-spheres

$$\rho^{(r)}(p, p') = \|p - \langle p, p'\rangle p'\| = \sqrt{1 - \langle p, p'\rangle^2} = \rho^{(r)}(p', p) \tag{2}$$

is a quasi-metric (symmetric, vanishing on the diagonal $p = p'$ and satisfying the triangle inequality). On general manifolds, however, the *residual distance* $\rho^{(r)}$ may be neither symmetric nor satisfying the triangle inequality.

Obviously, for X uniformly distributed on a sphere, the entire sphere is identical with the set of intrinsic, extrinsic and residual means: non-unique intrinsic and extrinsic means may depend counterintuitively on the dimension of the ambient space. Here is a simple illustration.

Proposition 1 *Suppose that X is a random point on a unit sphere S^{D-1} that is uniformly distributed on a unit subsphere S. Then*

(i) *every point on S^{D-1} is an extrinsic mean and,*
(ii) *if S is a proper subsphere then the set of intrinsic means is equal to the unit subsphere S' orthogonal to S.*

Proof The first assertion is a consequence of $\rho^{(e)}(x, y)^2 + \rho^{(e)}(x, -y)^2 = 4$ for every $x, y \in S^{D-1}$. The second assertion follows from

$$\begin{aligned}\rho^{(i)}(x, y)^2 + \rho^{(i)}(x, -y)^2 &= \rho^{(i)}(x, y)^2 + \left(\pi - \rho^{(i)}(x, y)\right)^2 \\ &\geq \frac{\pi^2}{2} \text{ for every } x, y \in S^{D-1}\end{aligned}$$

for the intrinsic distance $\rho^{(i)}(x, y) = 2 \arcsin(\|x - y\|/2)$ with equality if and only if x is orthogonal to y. □

Proposition 2 *If a random point X on a unit sphere is a.s. contained in a unit subsphere S then S contains every residual mean as well as every unique intrinsic or extrinsic mean.*

Proof Suppose that $x = v + \nu$ is a mean of X with $v/\|v\| \in S$ and $\nu \in S^{D-1}$ normal to S. Since $1 - \langle X, v + \nu\rangle^2 = 1 - \langle X, v\rangle^2 \geq 1 - \langle X, v\rangle^2/\|v\|^2$ a.s. with equality if and only if $\nu = 0$, the assertion for residual means follows at once from (2). For intrinsic and extrinsic means we argue with $\|X - (v + \nu)\| = \|X - (v - \nu)\|$ a.s. yielding $\nu = 0$ in case of uniqueness. □

Let us now incorporate more of the structure common to shape spaces. Expanding the definition due to (Kendall et al. 1999, p. 249), we start by assuming that the shape space is a quotient modulo a *proper group action* of a Lie group G on a manifold M, i.e. that every sequence $g_n \in G$ has a point of accumulation $g \in G$ if there are $p, p' \in M$ and a sequence $M \ni p_n \to p$ such that $g_n p_n \to p'$ (cf. Palais 1961). In consequence, the orbits $[p] = \{gp : g \in G\}$ are closed in M and the canonical quotient $Q = M/G$ is Hausdorff. Obviously, every compact group such as $SO(m)$ or $O(m)$ acts properly. Examples of non-compact groups acting properly can be found in projective shape analysis (cf. Kent et al. 2011).

Definition 3 A complete connected finite-dimensional Riemannian manifold M with geodesic distance d_M on which a Lie group G acts properly and isometrically from the left is called a *pre-shape space*. Moreover the canonical quotient

$$\pi : M \to Q := M/G = \{[p] : p \in M\} \text{ with the } \textit{orbit } [p] = \{gp : g \in G\},$$

is called a *shape space*.

As a consequence of the isometric action we have that $d_M(gp, p') = d(p, g^{-1}p')$ for all $p, p' \in M$, $g \in G$. For $p, p' \in M$ we say that p is in *optimal position* to p' if $d_M(p, p') = \min_{g\in G} d_M(gp, p')$, the minimum is attained since the action is proper. As is well known (e.g. Bredon 1972, p. 179) there is an open and dense submanifold M^* of M such that the canonical quotient $Q^* = M^*/G$ restricted to M^* carries a natural manifold structure also being open and dense in Q. Elements in M^* and Q^*, respectively, are called *regular*, the complementary elements are *singular*; Q^* is the *manifold part* of Q.

Intrinsic means on shape spaces. The canonical quotient distance

$$d_Q([p], [p']) := \min_{g\in G} d_M(gp, p') = \min_{g,h\in G} d_M(gp, hp')$$

is called *intrinsic distance* and the corresponding d_Q-Fréchet mean sets are called *intrinsic means*. Note that the intrinsic distance on Q^* is equal to the canonical geodesic distance.

Ziezold and Procrustean means. Now, assume that we have an embedding with orthogonal projection $\Phi : \mathbb{R}^D \to M \subset \mathbb{R}^D$ as above. If the action of G is isometric w.r.t. the extrinsic metric, i.e. if $\|gp - gp'\| = \|p - p'\|$ for all $p, p' \in M$ and $g \in G$ then call

$$\rho_Q^{(z)}([p], [p']) := \min_{g\in G} \|gp - p'\| \quad \text{and}$$
$$\rho_Q^{(p)}([p], [p']) := \min_{\substack{g \in G,\ gp \text{ in} \\ \text{opt. pos. to } p'}} \|d\Phi_{p'}(gp - p')\|$$

the *Ziezold distance* and the *Procrustean distance* on Q, respectively. Call the corresponding population and sample Fréchet $\rho_Q^{(z)}$-means, respectively, the sets of population and sample *Ziezold means*, respectively. Similarly, call the corresponding population and sample Fréchet $\rho_Q^{(p)}$-means, respectively, the sets of population and sample *Procrustean means*, respectively.

We say that *optimal positioning is invariant* if for all $p, p' \in M$ and $g^* \in G$,

$$d_M(g^* p, p') = \min_{g \in G} d_M(gp, p') \Leftrightarrow \|g^* p - p'\| = \min_{g \in G} \|gp - p'\|.$$

Remark 2 Indeed for $Q = \Sigma_m^k, R\Sigma_m^k$, optimal positioning is invariant (cf. Kendall et al. 1999, p. 206), Procrustean means coincide with means of *general Procrustes analysis* introduced by Gower (1975) and Ziezold means coincide with means as introduced by Ziezold (1994) for Σ_2^k. For Σ_m^k, these means have already been introduced in Sect. 2. Moreover for Σ_2^k, Procrustean means agree with extrinsic means with respect to the Veronese–Whitney embedding (cf. Bhattacharya and Patrangenaru 2003; Sect. 5).

As previously defined in Sect. 2, Procrustean means on Σ_m^k are also called *full Procrustes means* in the literature to distinguish them from *partial Procrustes means* on Kendall's *size-and-shape spaces* not further discussed here (e.g. Dryden and Mardia 1998). We only note that partial Procrustes means are identical to the respective intrinsic, Procrustean and Ziezold means which on Kendall's size-and-shape spaces, all agree with one another. Ziezold means on Kendall's shape spaces have also been studied as partial Procrustes means of unit size confiugrations by Kendall et al. (1999).

4 Horizontal lifting and manifold stability

In this section we derive a measurable horizontal lifting and the stability theorem underlying Theorem 2. To this end we first recall how a shape space is made up from manifold strata of varying dimensions. Unless otherwise referenced, we use basic terminology that can be found in any standard textbook on differential geometry, e.g. Kobayashi and Nomizu (1963, 1969).

4.1 Preliminaries

Assume that $Q = M/G$ is a shape space as in Definition 3. $T_p M$ is the tangent space of M at $p \in M$ and $\exp_p$ denotes the *Riemannian exponential* at p. Recall that on a Riemannian manifold the *cut locus* $C(p)$ of p comprises all points q such that the extension of a length minimizing geodesic joining p with q is no longer minimizing beyond q. In consequence, on a complete and connected manifold M we have for every $p' \in M$ that there is $v' \in T_p M$ such that $p' = \exp_p v'$ while $v' = \exp_p^{-1} p'$ of minimal modulus is uniquely determined as long as $p' \in M \setminus C(p)$. It is well known that the cut locus has measure zero in the sense that its image in any local chart has Lebesgue measure zero. From now on we call the cut locus the *manifold cut locus* in order to distinguish it from the *quotient cut locus* $C^{\text{quot}}(q)$ of $q \in Q$ which we define as $C^{\text{quot}}(q) := \{[p'] : p' \in C(p) \text{ is in optimal position to some } p \in q\}$. Due to the isometric action we have for any $p \in q$ that

$$C^{\text{quot}}(q) = \{[p'] : p' \in C(p) \text{ is in optimal position to } p\} \subset \pi\big(C(p)\big). \tag{3}$$

Recall from Sect. 3 that Q contains an open and dense manifold part Q^*. Thus, for $q \in Q^*$ we can consider the quotient cut locus and the manifold cut locus, both of which are subsets of Q, in general different as the following Lemma teaches. In particular, quotient cut loci are void in the special case of Kendall's shape spaces.

Lemma 1 *$C(q) \neq \emptyset$ for every $q \in \Sigma_2^k$ while $C^{\text{quot}}(q) = \emptyset$ for all $q \in \Sigma_m^k$. Similarly $C^{\text{quot}}(q) = \emptyset$ for all $q \in R\Sigma_m^k$.*

Proof The first assertion follows from the fact that Σ_2^k is a compact manifold. For the second assertion consider $[p] \in \Sigma_m^k$. Since $C(p) = \{-p\}$ for $p \in S_m^k$ and $[p] = [-p]$ for even m as well as for odd m if p is not regular, and, since $p, -p$ are not in optimal position for odd m if p is regular, we have that $C^{\text{quot}}([p]) = \emptyset$. The third assertion follows from the fact that $[p] = [-p]$ for all $[p] \in R\Sigma_m^k$. □

Next we collect consequences of the isometric Lie group action, see Bredon (1972).

(A) With the *isotropy group* $I_p = \{g \in G : gp = p\}$ for $p \in M$, every orbit carries the natural structure of a coset space $[p] \cong G/I_p$. Moreover, $p' \in M$ is of *orbit type* (G/I_p) if $I_{p'} = gI_pg^{-1} = I_{gp}$ for a suitable $g \in G$. If $I_p \subset I_{gp'}$ for suitable $g \in G$ then p' is of *lower orbit type* than p and p is of *higher orbit type* than p'.

(B) The pre-shapes of equal orbit type $M^{(I_p)} := \{p' \in M : p' \text{ is of orbit type } (G/I_p)\}$ and the corresponding shapes $Q^{(I_p)} := \{[p'] : p' \in M^{(I_p)}\}$ are manifolds in M and Q, respectively. Moreover, for $q \in Q$ denote by $Q^{(q)}$ the shapes of higher orbit type.

(C) The orthogonal complement H_pM in T_pM of the tangent space $T_p[p]$ along the orbit is called the *horizontal space*: $T_pM = T_p[p] \oplus H_pM$.

(D) The *Slice Theorem* states that every $p \in M$ has a *tubular neighborhood* $[p] \subset U \subset M$ such that with a suitable subset $D \subset H_pM$ the *twisted product* $\exp_p D \times_{I_p} G$ is diffeomorphic with U. Here, the twisted product is the natural topological quotient of the product space $\exp_p D \times G$ modulo the equivalence

$$(\exp_p v, g) \sim_{I_p} (\exp_p v', g') \;\Leftrightarrow \exists h \in I_p \text{ such that } v' = dhv,\; g' = gh^{-1}.$$

We then say that the tubular neighborhood U *admits a slice* $\exp_p D$ *via* $U \cong \exp_p D \times_{I_p} G$.

(E) Every $p \in M$ has a tubular neighborhood U of p that admits a slice $\exp_p D$ such that every $p' \in \exp_p D$ is in optimal position to p. Moreover, for any tubular neighborhood U admitting a slice $\exp_p D$, all points $p' \in U$ are of orbit type higher than or equal to the orbit type of p and only finitely many orbit types occur in U. If p is regular, i.e. of maximal orbit type, then the product is trivial: $\exp_p D \times_{I_p} G \cong \exp_p D \times G/I_p$.

Finally let us extend the following uniqueness property for the intrinsic distance to the Ziezold distance. The differential of the mapping $f_{\text{int}}^{p'}: M \setminus C(p') \to [0,\infty)$ defined by $f_{\text{int}}^{p'}(p) = d_M(p, \exp_p p')^2$ is given by $df_{\text{int}}^{p'}(p) = -2v$ with $v = \exp_p^{-1} p'$ (cf. Kobayashi and Nomizu 1969, p. 110; Karcher 1977). Hence, we have for $p_1, p_2 \in M \setminus C(p)$ that

$$df_{\text{int}}^{p_1}(p) = df_{\text{int}}^{p_2}(p) \;\Leftrightarrow\; p_1 = p_2. \tag{4}$$

In view of the extrinsic distance let $f_{\text{ext}}^{p'}: M \setminus C(p') \to [0,\infty)$ be defined by $f_{\text{ext}}^{p'}(p) = \|p - p'\|^2 = \|p - \exp_p(\exp_p^{-1} p')\|^2$. Mimicking (4) introduce the following condition

$$df_{\text{ext}}^{p_1}(p) = df_{\text{ext}}^{p_2}(p) \Leftrightarrow p_1 = p_2 \tag{5}$$

for $p_1, p_2 \in M \setminus C(p)$.

Remark 3 (5) is valid on closed half spheres since on the unit sphere

$$df_{\text{ext}}^{p'}(p) = -2\,\frac{v}{\|v\|}\,\sin(\|v\|) \text{ with } v = \exp_p^{-1} p'.$$

4.2 A measurable horizontal lift

To establish the stability of means in Theorem 5 in the following Sect. 4.3, here we lift a random shape X from Q horizontally to a random pre-shape Y on M. In order to do so we need to guarantee the measurability of the horizontal lift in Theorem 3 below, the proof of which can be found in the Appendix.

Before continuing, let us consider a simple example for illustration. Suppose that $G = S^1 \subset \mathbb{C}$ acts on $M = \mathbb{C}$ by complex scalar multiplication. Then $[0,\infty) \cong Q = M/G$ having the two orbit types $(S^1/I_0) = \{1\}$ and $(S^1/I_1) = S^1$ gives rise to $Q^{(I_0)} = \{0\}$ and $Q^{(I_1)} = (0,\infty)$. Obviously, M admits a global slice via the polar decomposition $[0,\infty) \times_{S^1} S^1 = \{0\} \cup \big((0,\infty) \times S^1\big) \cong M$ about $0 \in M$ (the Riemannian exponential is the identity if $T_0\mathbb{C}$ is identified with $\mathbb{C}$). Here, X can be identified with its horizontal lift Y to the global slice $[0,\infty) \subset M$. If, say, X is uniformly distributed on $[1,2]$ then $\mathbb{P}\{X \in Q^{(I_1)}\} > 0$. In this case the stability theorem states the obvious fact that $0 \in Q^{(I_0)}$ cannot be a mean of X.

Definition 4 Call a measurable subset $L \subset M$ a *measurable horizontal lift* of a measurable subset R of M/G *in optimal position to* $p \in M$ if

the canonical projection $L \to R \subset M/G$ is surjective,
every $p' \in L$ is in optimal position to p,
every orbit $[p']$ of $p' \in L$ meets L once.

Theorem 3 *Let $p \in [p] \in Q$ and $A \subset Q$ countable. Then there is a measurable horizontal lift L of $Q^{([p])} \cup A$ in optimal position to p.*

Theorem 4 *Assume that X is a random shape on Q and that there are $p \in M$ and $A \subset Q$ countable such that X is supported by $\left(Q^{([p])} \cup A\right) \setminus C^{\text{quot}}([p])$. With a measurable horizontal lift L of $\left(Q^{([p])} \cup A\right) \setminus C^{\text{quot}}([p])$ in optimal position to p define the random element Y on $L \subset M$ by $\pi \circ Y = X$.*

(i) *If $[p]$ is an intrinsic mean of X on Q, then p is an intrinsic mean of Y on M and*

$$\mathbb{E}(\exp_p^{-1} Y) = 0.$$

(ii) *If $[p]$ is a Ziezold mean of X on Q and optimal positioning is invariant, then p is an extrinsic mean of Y on M and*

$$\mathbb{E}\big(df_{\text{ext}}^{Y}(p)\big) = 0.$$

(iii) *If $[p]$ is a Procrustean mean of X on Q and optimal positioning is invariant, then p is a residual mean of Y on M.*

Proof Suppose that $[p]$ is an intrinsic mean of X. If p were not an intrinsic mean of Y, there would be some $M \ni p' \neq p$ leading to the contradiction

$$\mathbb{E}\big(d_Q([p'], X)^2\big) = \mathbb{E}\big(d_M(p', Y)^2\big) \;<\; \mathbb{E}\big(d_M(p, Y)^2\big) \;=\; \mathbb{E}\big(d_Q([p], X)^2\big).$$

Hence, p is an intrinsic mean of Y. Replacing d_Q by $\rho_Q^{(z)}$ and d_M by the Euclidean distance gives the assertion for Ziezold and extrinsic means, respectively; and, using the Procrustean distance $\rho_Q^{(p)}$ on Q as well as the residual distance $\rho_M^{(r)}$ on M gives the assertion for Procrustean and residual means, respectively.

For intrinsic means $p \in M$, the necessary condition $\mathbb{E}\big(\exp_p^{-1} Y\big) = 0$ is developed in Kobayashi and Nomizu (1969, p. 110), cf. also Karcher (1977) and Kendall (1990, p. 395), which yields the asserted equality in (i). By definition, the analog condition for an extrinsic mean $p \in M$ is $\mathbb{E}\big(df_{\text{ext}}^{Y}(p)\big) = 0$ which is the asserted equality in (ii) completing the proof. □

Remark 4 Since the maximal intrinsic distance on Σ_m^k and $R\Sigma_m^k$ is

$$\frac{\pi}{2} = \max_{x,y \in S_m^k} \min_{g \in SO(m)} \arccos\left(\operatorname{tr}(gxy^T)\right) = \max_{x,y \in S_m^k} \min_{g \in O(m)} \arccos\left(\operatorname{tr}(gxy^T)\right),$$

taking into account Remark 3, Condition (5) is satisfied for any horizontal lift in optimal position.

4.3 Manifold stability

The proof of the following central theorem is deferred to the Appendix.

Theorem 5 *Assume that X is a random shape on Q, $p \in M$ and that $A \subset Q$ is countable such that X is supported by $\big(Q^{([p])} \cup A\big) \setminus C^{\text{quot}}([p])$ and let $p' \in [p'] \in Q^{([p])}$. If $\mathbb{P}\{X \in Q^{(I_{p'})}\} \neq 0$ and if either $[p]$ is*

(i) *an intrinsic mean of X or*
(ii) *a Ziezold mean of X while optimal positioning is invariant and (5) is valid,*

then p' is of lower orbit type than p.

We have at once the following Corollary.

Corollary 1 (Manifold stability theorem) *Suppose that X is a random shape on Q that is supported by $Q \setminus C^{\text{quot}}([p])$ for some $[p] \in Q$ assuming the manifold part Q^* with non-zero probability and having at most countably many point masses on the singular part. Then $[p]$ is regular if it is an intrinsic mean of X, or if it is a Ziezold mean, optimal positioning is invariant and (5) is valid.*

Since $Q \setminus Q^{(q)}$ is a null set in Q for every $q \in Q$ (cf. Bredon 1972, p. 184) and so is $C^{\text{quot}}(q)$—by (3) it is contained in the projection of a null set—we have the following practical application.

Corollary 2 *Suppose that a random shape on Q is absolutely continuously distributed with respect to the projection of the Riemannian volume on M. Then intrinsic and Ziezold population means are regular; the latter if optimal positioning is invariant and (5) is valid. In addition, intrinsic and Ziezold sample means are a.s. regular.*

4.4 An example for non-stability of Procrustean means

Consider a random configuration $Z \in F_3^4$ assuming the collinear quadrangle q_1 with probability $2/3$ and the planar quadrangle q_2 with probability $1/3$ where

$$q_1 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad q_2 = \begin{pmatrix} 1 & 1 & -2 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & -\frac{3}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Corresponding pre-shapes in optimal position w.r.t. the action of $SO(3)$ and $O(3)$ are given by

$$p_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that $[p_2]$ has regular shape in $(\Sigma_3^4)^*$. The full Procrustes mean of $[Z] \in \Sigma_3^4$ is easily computed to have the singular shape $[p_1] \in \Sigma_3^4 \setminus (\Sigma_3^4)^*$, see Fig. 2 as well as Examples 1 and Sect. 7.2, cf. also Remark 1.

5 Extrinsic means for Kendall's (reflection) shape spaces

Let us recall the well known *Veronese–Whitney* embedding for Kendall's planar shape spaces Σ_2^k. Identify F_2^k with $\mathbb{C}^{k-1}\setminus\{0\}$ such that every landmark column corresponds to a complex number. This means in particular that $z \in \mathbb{C}^{k-1}$ is a complex row(!)-vector. With the Hermitian conjugate $a^* = (\overline{a_{kj}})$ of a complex matrix $a = (a_{jk})$ the pre-shape sphere S_2^k is identified with $\{z \in \mathbb{C}^{k-1} : zz^* = 1\}$ on which $SO(2)$ identified with $S^1 = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ acts by complex scalar multiplication. Then, the well-known Hopf-Fibration mapping to complex projective space gives $\Sigma_2^k = S_2^k/S^1 = \mathbb{C}P^{k-2}$. Moreover, denoting with $M(k-1, k-1, \mathbb{C})$ all complex $(k-1) \times (k-1)$ matrices, the Veronese–Whitney embedding is given by

$$S_2^k/S^1 \to \{a \in M(k-1, k-1, \mathbb{C}) : a^* = a\}, \quad [z] \mapsto z^*z.$$

Remark 5 The Procrustean metric of Σ_2^k is isometric with the Euclidean metric of $M(k-1, k-1, \mathbb{C})$ since we have $\langle z, w\rangle = \mathrm{Re}(zw^*)$ for $z, w \in S_2^k$ and hence, $d^{(p)}_{\Sigma_2^k}([z], [w]) = \sqrt{1 - wz^*zw^*} = \|w^*w - z^*z\|/\sqrt{2}$.

The idea of the Veronese–Whitney embedding can be carried to the general case of shapes of arbitrary dimension $m \geq 2$. Even though the embedding given below is apt only for reflection shape space it can be applied to practical situations in similarity shape analysis whenever the geometrical objects considered have a common orientation. As above, the number k of landmarks is essential and will be considered fixed throughout this section; the dimension $1 \leq m < k$, however, is lost in the embedding and needs to be retrieved by projection. To this end recall the embedding of S_j^k in S_m^k $(1 \leq j \leq m)$ in (1) which gives rise to a canonical embedding of $R\Sigma_j^m$ in $R\Sigma_m^k$. Moreover, consider the strata

$$(R\Sigma_m^k)^j := \{[x] \in R\Sigma_m^k : \mathrm{rank}(x) = j\}, \quad (\Sigma_m^k)^j := \{[x] \in \Sigma_m^k : \mathrm{rank}(x) = j\}$$

for $j = 1, \ldots, m$, each of which carries a canonical manifold structure; due to the above embedding, $(R\Sigma_m^k)^j$ will be identified with $(R\Sigma_j^k)^j$ such that

$$R\Sigma_m^k = \bigcup_{j=1}^m (R\Sigma_j^k)^j ,$$

and $(R\Sigma_m^k)^j$ with $(\Sigma_m^k)^j$ in case of $j < m$. At this point we note that $SO(m)$ is connected, while $O(m)$ is not; and the consequences for the respective manifold parts, i.e. points of maximal orbit type:

$$(\Sigma_m^k)^* = (\Sigma_m^k)^{m-1} \cup (\Sigma_m^k)^m , \quad (R\Sigma_m^k)^* = (R\Sigma_m^k)^m. \tag{6}$$

Similarly, we have a stratifiction

$$\mathcal{P} := \left\{a \in M(k-1,k-1) : a = a^T \geq 0, \operatorname{tr}(a) = 1\right\} = \bigcup_{j=1}^{k-1} \mathcal{P}^j$$

of a compact flat convex space $\mathcal{P}$ with non-flat manifolds $\mathcal{P}^j := \{a \in \mathcal{P} : \operatorname{rank}(a) = j\}$ $(j = 1, \dots, k-1)$, all embedded in $M(k-1, k-1)$. The *Schoenberg map* $\mathfrak{s} : R\Sigma_m^k \to \mathcal{P}$ is then defined on each stratum by

$$\mathfrak{s}|_{(R\Sigma_m^k)^j} =: \mathfrak{s}^j : (R\Sigma_m^k)^j \to \mathcal{P}^j\,, \quad [x] \mapsto x^T x\,.$$

For $x \in S_j^k$ recall the tangent space decomposition $T_x S_j^k = T_x[x] \oplus H_x S_j^k$ into the *vertical tangent space* along the orbit $[x]$ and its orthogonal complement the *horizontal tangent space*. For $x \in (S_j^k)^j$ identify canonically (cf. Kendall et al. 1999, p. 109):

$$T_{[x]}(R\Sigma_j^k)^j \cong H_x S_j^k = \{w \in M(j, k-1) : \operatorname{tr}(wx^T) = 0, wx^T = xw^T\}.$$

Then the assertion of the following Theorem condenses results of Bandulasiri and Patrangenaru (2005), cf. also Dryden et al. (2008).

Theorem 6 *Each $\mathfrak{s}^j$ is a diffeomorphism with inverse $(\mathfrak{s}^j)^{-1}(a) = [(\sqrt{\lambda}u^T)_1^j]$ where $a = u\lambda u^T$ with $u \in O(k-1)$, $\lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$, and $0 = \lambda_{j+1} = \dots = \lambda_{k-1}$ in case of $j < k-1$. Here, $(a)_1^j$ denotes the matrix obtained from taking only the first j rows from a. For $x \in S_j^k$ and $w \in H_x S_j^k \cong T_{[x]}(R\Sigma_j^k)^j$ the derivative is given by*

$$d(\mathfrak{s}^j)_{[x]} w = x^T w + w^T x.$$

Remark 6 In contrast to the Veronese-Whitney embedding, the Schoenberg embedding is not isometric as the example of

$$x = \begin{pmatrix} \cos\phi & 0 \\ 0 & \sin\phi \end{pmatrix}, \quad w_1 = \begin{pmatrix} \sin\phi & 0 \\ 0 & -\cos\phi \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 & \cos\phi \\ \sin\phi & 0 \end{pmatrix},$$

teaches: $\|x^T w_1 + w_1^T x\| = \sqrt{2}\,2|\cos\phi\sin\phi|$, $\|x^T w_2 + w_2^T x\| = \sqrt{2}$.

Since $\mathcal{P}$ is bounded, convex and Euclidean, the classical expectation $\mathbb{E}(X^T X) \in \mathcal{P}^j$ for some $1 \leq j \leq k-1$ of the Schoenberg image $X^T X$ of an arbitrary random reflection shape $[X] \in R\Sigma_m^k$ is well defined. Then we have at once the following relation between the rank of the Euclidean mean and increasing sample size.

Theorem 7 *Suppose that a random reflection shape $[X] \in R\Sigma_m^k$ is distributed absolutely continuous w.r.t. the projection of the spherical volume on S_m^k. Then*

$$\mathbb{E}(X^T X) \in \mathcal{P}^{k-1} \ \textit{and} \ \ \frac{1}{n}\sum_{i=1}^n X_i^T X_i \in \mathcal{P}^{\min(nm,k-1)} \ a.s.$$

for every i.i.d. sample $X_1, \dots, X_n \sim X$.

Hence, in statistical settings involving a higher number of landmarks, a sufficiently well behaved projection of a high rank Euclidean mean onto lower rank $\mathcal{P}^r \cong (\Sigma_r^k)^r$, usually $m = r$, is to be employed, giving at once a mean shape satisfying strong consistency and a CLT. Here, unlike to intrinsic or Procrustes analysis, the dimension r chosen is crucial for the dimensionality of the mean obtained.

The orthogonal projection

$$\phi^r : \bigcup_{i=r}^{k-1} \mathcal{P}^i \to \mathcal{P}^r \,, \quad a \;\mapsto\; \operatorname{argmin}_{b\in\mathcal{P}^r} \operatorname{tr}\left((a-b)^2\right)$$

giving the set of *extrinsic Schoenberg means* has been computed by Bhattacharya (2008):

Theorem 8 *For* $1 \leq r \leq k-1$, $a = u\lambda u^T \in \mathcal{P}$ *with* $u \in O(k-1)$, $\lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$, $\lambda_1 \geq \cdots \geq \lambda_{k-1}$ *and* $\lambda_r > 0$ *the orthogonal projection onto* $\mathcal{P}^r$ *is given by* $\phi^r(a) = u\mu u^T$ *with* $\mu = \operatorname{diag}(\mu_1, \ldots, \mu_r, 0, \ldots, 0)$,

$$\mu_i = \lambda_i + \frac{1}{r} - \overline{\lambda}_r \; (i = 1, \ldots, r)$$

and $\overline{\lambda}_r = \frac{1}{r}\sum_{i=1}^r \lambda_i \leq \frac{1}{r}$ *which is uniquely determined if and only if* $\lambda_r > \lambda_{r+1}$.

With the notation of Theorem 8, a non-orthogonal *central projection* $\psi^r(a) = u\nu u^T$ equally well and uniquely determined has been proposed by Dryden et al. (2008) with

$$\nu = \operatorname{diag}(\nu_1, \ldots, \nu_r, 0, \ldots, 0), \quad \nu_i = \frac{\lambda_i}{r\overline{\lambda}_r} \; (i = 1, \ldots, r).$$

Orthogonal and central projection are depicted in Fig. 1.

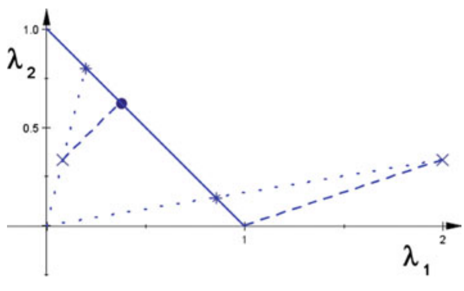

Fig. 1 Projections (if existent) of two points (*crosses*) in the λ-plane to the open line segment $\Lambda = \{(\lambda_1, \lambda_2) : \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 > 0\}$. The *dotted line* gives the central projections (*denoted by stars*) which is well defined for all symmetric, positive definite matrices (corresponding to the first open quadrant), the *dashed line* gives the orthogonal projection (*circle*) which is well defined in the *triangle* below Λ (corresponding to $\mathcal{P}$) and above Λ in an *open strip*. In particular, it exists not for the right point

6 Testing, local effects of curvature and loci

In this section we assume that a manifold stratum M supporting a random element X is isometrically embedded in a Euclidean space $\mathbb{R}^D$ of dimension $D > 0$. With the orthogonal projection $\Phi : \mathbb{R}^D \to M$ from Sect. 3 and the Riemannian exponential $\exp_p$ of M at p we have the

intrinsic tangent space coordinate	$\exp_p^{-1} X$ and the
residual tangent space coordinate	$d\Phi_p(X - p)$,

respectively, of X at p, if existent.

6.1 The two sample test for equality of means

Let ϕ_p denote a local chart of M near p. E.g. ϕ_p can be one of the above tangent space coordinates. The following Theorem is taken from Huckemann (2011), cf. also Jupp (1988), Hendriks and Landsman (1996), and Bhattacharya and Patrangenaru (2005).

Theorem 9 (CLT for Fréchet ρ-means) *Let μ be a unique Fréchet ρ-mean of a random M-valued variable X. If $X \to \rho^2(X, \mu)$ is twice continuously differentiable on the support of X, if first moments of the second derivatives and the second moments of the first derivatives of $p \to \rho(X, p)^2$ exist and are continuous near $p = \mu$ and if μ_n is a measurable selection of the sample mean set, then there are symmetric semi-positive definite matrices A, Σ such that*

$$\sqrt{n} A\phi_\mu(\mu_n) \to \mathcal{N}(0, \Sigma) \text{ in distribution as } n \to \infty.$$

For the following suppose that $X_1, \ldots, X_n \overset{i.i.d}{\sim} X$ and $Y_1, \ldots, Y_m \overset{i.i.d}{\sim} Y$ are independent samples on M with unique Fréchet ρ-means μ_X and μ_Y, respectively ($m, n > 0$). Moreover assume that under the null hypothesis $\mu_X = \mu_Y = \mu$ the support of X and Y is contained in the domain of definition of ϕ_μ. Then under the hypotheses of Theorem 9, if A has maximal rank (which is the case in most realistic situations) the classical Hotelling T^2 statistic $T^2(n, m)$ of $\phi_{\mu_{n+m}}(X_1), \ldots, \phi_{\mu_{n+m}}(X_n)$ and $\phi_{\mu_{n+m}}(Y_1), \ldots, \phi_{\mu_{n+m}}(Y_m)$ is well defined for $n+m$ sufficiently large where μ_{n+m} denotes a measurable selection of a pooled Fréchet ρ-sample mean (e.g. Anderson 2003, Chapter 5). Finally, we assume that there is a constant $C > 0$ such that $\|\phi_p(X) - \phi_\mu(X)\|, \|\phi_p(Y) - \phi_\mu(Y)\| \le C\|p - \mu\|$ a.s. for p near μ. This latter condition is fulfilled for intrinsic and extrinsic coordinates if X and Y have compact support.

Theorem 10 (Two-sample test) *Under the above hypotheses for $n, m \to \infty$, $T^2(n, m)$ is asymptotically Hotelling T^2-distributed if either $n/m \to 1$ or if* $\mathrm{COV}(\phi_\mu(X)) = \mathrm{COV}(\phi_\mu(Y))$.

Proof Let $P_1, \ldots, P_{n+m}$ for $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ and consider the Euclidean data $Z_j^{(n,m)} = \phi_{\mu_{n+m}}(P_j)$, $Z_j = \phi_\mu(P_j)$ for $j = 1, \ldots, n + m$, $n, m \in \mathbb{N}$. Since the Z_j are independent (Lehmann 1997, p. 462) yields that the corresponding Hotelling

T^2-statistic is asymptotically Hotelling T^2 distributed under either condition of the Theorem. Since $Z_j^{(n,m)} = Z_j + O_p\big(1/\sqrt{n}\big)$ by Theorem 9 and the uniform Lipschitz condition on $p \to \phi_p(X)$, the same holds for the Hotelling T^2 statistic $T(n,m)$ obtained from the $Z_j^{(n,m)}$. □

The hypotheses of Theorems 9 and 10 are fulfilled if X, Y have compact support not containing the cut locus $C(\mu)$ and if ϕ is an intrinsic or extrinsic tangent space coordinate.

6.2 Finite power of tests and tangent space coordinates

With the above setup, assume that $\mu \in M$ is a unique mean of X. Moreover, we assume that M is curved near μ, i.e. that there is $c \in \mathbb{R}^D$, the center of the osculatory circle touching the geodesic segment in M from X to μ at μ with radius r. If X_r is the orthogonal projection of X to that circle, then $X = X_r + O(\|X-\mu\|^3)$. Moreover, with

$$\cos\alpha = \left\langle \frac{X-c}{\|X-c\|}, \frac{\mu-c}{r} \right\rangle \;=\; \frac{1}{r^2}\,\langle X_r - c, \mu - c\rangle + O(\|X-\mu\|^3)$$

we have the residual tangent space coordinate

$$v = X - c - \frac{\mu - c}{r}\,\|X-c\|\,\cos\alpha = X_r - c - (\mu - c)\cos\alpha + O(\|X-\mu\|^3)$$

having squared length $\|v\|^2 = r^2 \sin\alpha^2 + O(\|X-\mu\|^3)$. By isometry of the embedding, the intrinsic tangent space coordinate is given by

$$\exp_\mu^{-1} X = \frac{r\alpha}{\|v\|}\, v + O(\|X-\mu\|^3).$$

With the component

$$\nu = \mu - c - \|X-c\|\frac{\mu-c}{r}\,\cos\alpha = (\mu - c)(1-\cos\alpha) + O(\|X-\mu\|^3)$$

of X normal to the above mentioned geodesic segment of squared length $\|\nu\|^2 = r^2(1-\cos\alpha)^2 + O(\|X-\mu\|^3)$, we obtain $\|\exp_\mu^{-1} X\|^2 \;=\; \|v\|^2 + \|\nu\|^2 + O(\|X-\mu\|^3)$, since

$$(1-\cos\alpha)^2 + \sin^2\alpha = 2(1-\cos\alpha) \;=\; \alpha^2 + 2\frac{\alpha^4}{4!} + \cdots$$

and $\alpha = O(\|X-\mu\|)$. In consequence we have

Remark 7 In approximation, the variation of intrinsic tangent space coordinates is the sum of the variation $\|v\|^2$ of residual tangent space coordinates and the variation normal to it. In particular due to Pythagoras' theorem,

$$\| \exp_\mu^{-1} X\|^2 \;\geq\; \|v\|^2 + \|\nu\|^2$$

whenver the l.h.s. is defined. Since for a two-sample test for equality of means (cf. Sect. 6.1) even under the alternative, the means in the normal coordinate tend to agree with one another (especially for $n = m$), a higher power for tests based on intrinsic means can be expected when solely residual tangent space coordinates obtained from an isometric embedding are used rather than intrinsic tangent space coordinates. This effect, however, is only of order $n^{-\frac{3}{2}}$ (cf. Hendriks and Landsman 1996).

Note that the natural tangent space coordinates for Ziezold means are residual.

A simulated classification example in Sect. 7 illustrates this effect.

6.3 The 1:3-property for spherical and Kendall shape means

In this section $M = S^{D-1} \subset \mathbb{R}^D$ is the $(D-1)$-dimensional unit-hypersphere embedded isometrically in Euclidean D-dimensional space. The orthogonal projection $\Phi : \mathbb{R}^D \to S^{D-1} : p \to \frac{p}{\|p\|}$ is well defined except for the origin $p = 0$, and the normal space at $p \in S^{D-1}$ is spanned by p itself. In consequence, a random point X on S^{D-1} has

$$d\Phi_p(X-p) = X - p\cos\alpha, \quad \exp_p^{-1}(X) = \begin{cases} \frac{\alpha}{\sin\alpha}\, d\Phi_p(X-p) & \text{for } X \neq p \\ 0 & \text{for } X = p \end{cases}$$

as residual and intrinsic, resp., tangent space coordinate at $-X \neq p \in S^{D-1}$ where $\cos\alpha = \langle X, p\rangle$, $\alpha \in [0,\pi)$.

Theorem 11 *If X a.s. is contained in an open half sphere, it has a unique intrinsic mean which is assumed in the interior of that half sphere.*

Proof Below we show that every intrinsic mean necessarily lies within the interior of the half sphere. Then, Kendall (1990, Theorem 7.3) yields uniqueness. W.l.o.g. let $X = (\sin\phi, x_2, \ldots, x_n)$ such that $\mathbb{P}\{\sin\phi \leq 0\} = 0 = 1 - \mathbb{P}\{\sin\phi > 0\}$ and assume that $p = (\sin\psi, p_2, \ldots, p_n) \in S^{D-1}$ is an intrinsic mean, $-\pi/2 \leq \phi, \psi \leq \pi/2$. Moreover let $p' = (\sin(|\psi|), p_2, \ldots, p_n)$. Since

$$\begin{aligned}
\mathbb{E}\left(\| \exp_p^{-1}(X)\|^2\right) &= \mathbb{E}\left(\arccos^2\langle p, X\rangle\right) \\
&= \mathbb{E}\left(\arccos^2\left(\sin\psi\sin\phi + \sum_{j=2}^n p_j x_j\right)\right) \\
&\geq \mathbb{E}\left(\| \exp_{p'}^{-1}(X)\|^2\right)
\end{aligned}$$

with equality if and only if $\sin|\psi| = \sin\psi$, this can only happen for $\sin\psi \geq 0$. Now, suppose that $p = (0, p_2, \ldots, p_n)$ is an intrinsic mean. For small deterministic $\psi \geq 0$ consider $p(\psi) = (\sin\psi, p_1 \cos\psi, \ldots, p_n \cos\psi)$. Then

$$\begin{aligned}\mathbb{E}\left(\| \exp_{p(\psi)}^{-1}(X)\|^2\right) &= \mathbb{E}\left(\arccos^2\left(\sin\psi \sin\phi + \cos\psi \sum_{j=2}^{n} p_j x_j\right)\right)\\ &= \mathbb{E}\left(\| \exp_p^{-1}(X)\|^2\right) - C_1\psi + O(\psi^2)\end{aligned}$$

with $C_1 > 0$ since $\mathbb{P}\{\sin\phi > 0\} > 0$. In consequence, p cannot be an intrinsic mean. Hence, we have shown that every intrinsic mean is contained in the interior of the half sphere. □

Remark 8 For the special case of spheres, this is a simple proof for the general theorem recently established by Afsari (2011) which extends results of Karcher (1977), Kendall (1990) and Le (2001, 2004), stating that the intrinsic mean on a general manifold is unique if among others the support of the distribution is contained in a geodesic half ball.

The following theorem characterizes the three spherical means.

Theorem 12 *Let X be a random point on S^{D-1}. Then $x^{(e)} \in S^{D-1}$ is the unique extrinsic mean if and only if the Euclidean mean $\mathbb{E}(X) = \int_{S^{D-1}} X \, d\,\mathbb{P}_X$ is non-zero. In that case*

$$\lambda^{(e)} x^{(e)} = \mathbb{E}(X)$$

with $\lambda^{(e)} = \|\mathbb{E}(X)\| > 0$. Moreover, there are suitable $\lambda^{(r)} > 0$ and $\lambda^{(i)} > 0$ such that every residual mean $x^{(r)} \in S^{D-1}$ satisfies

$$\lambda^{(r)} x^{(r)} = \mathbb{E}\big(\langle X, x^{(r)}\rangle \, X\big)\,,$$

and every intrinsic mean $x^{(i)} \in S^{D-1}$ satisfies

$$\lambda^{(i)} x^{(i)} = \mathbb{E}\left(\frac{\arccos\langle X, x^{(i)}\rangle}{\sqrt{1 - \langle X, x^{(i)}\rangle^2}} \, X\right).$$

In the last case we additionally require that $\mathbb{E}\left(\frac{\arccos\langle X, x^{(i)}\rangle}{\sqrt{1-\langle X, x^{(i)}\rangle^2}} \langle X, x^{(i)}\rangle\right) > 0$ which is in particular the case if X is a.s. contained in an open half sphere.

Proof The assertions for the extrinsic mean are well known from Hendriks et al. (1996). The second assertion for residual means follows from minimization of

$$\int_{S^{D-1}} \|p - \langle p, x\rangle x\|^2 \, d\,\mathbb{P}_X(p) = 1 - \int_{S^{D-1}} \langle p, x\rangle^2 \, d\,\mathbb{P}_X(p)$$

with respect to $x \in \mathbb{R}^D$ under the constraining condition $\|x\| = 1$. Using a Lagrange ansatz this leads to the necessary condition

$$\int_{S^{D-1}} \langle p, x \rangle \, p \, d\, \mathbb{P}_X(p) = \lambda x$$

with a Lagrange multiplier λ of value $\mathbb{E}(\langle X, x\rangle^2)$ which is positive unless X is supported by the hypersphere orthogonal to x. In that case, by Proposition 2, x cannot be a residual mean of X, as every residual mean is as well contained in that hypersphere. Hence, we have $\lambda^{(r)} := \lambda > 0$.

The Lagrange method applied to

$$\int_{S^{D-1}} \| \exp_x^{-1}(p)\|^2 \, d\, \mathbb{P}_X(p) = \int_{S^{D-1}} \arccos^2(\langle p, x\rangle) \, d\, \mathbb{P}_X(p)$$

taking into account Theorem 11, insuring that $x^{(i)}$ is in the open half sphere that contatins X a.s., yields the third assertion on the intrinsic mean. □

Recall that residual means are eigenvectors to the largest eigenvalue of the matrix $\mathbb{E}(XX^T)$. As such, they rather reflect the mode than the classical mean of a distribution:

Example 1 Consider $\gamma \in (0, \pi)$ and a random variable X on the unit circle $\{e^{i\theta} : \theta \in [0, 2\pi)\}$ which takes the value 1 with probability $2/3$ and $e^{i\gamma}$ with probability $1/3$. Then, explicit computation gives the unique intrinsic and extrinsic mean as well as the two residual means

$$x^{(i)} = e^{i\frac{\gamma}{3}}, \quad x^{(e)} = e^{i \arctan \frac{\sin\gamma}{2+\cos\gamma}}, \quad x^{(r)} = \pm\, e^{i\frac{1}{2}\arctan\frac{\sin(2\gamma)}{2+\cos(2\gamma)}}.$$

Figure 2 shows the case $\gamma = \frac{\pi}{2}$.

In contrast to Fig. 2, one may assume in many practical applications that the mutual distances of the unique intrinsic mean $x^{(i)}$, the unique extrinsic mean $x^{(e)}$ and the unique residual mean $x^{(r_0)}$ closer to $x^{(e)}$ are rather small, namely of the same order as the squared proximity of the modulus $\|\mathbb{E}(X)\|$ of the Euclidean mean to 1 (cf. Table 2). We will use the following condition

$$\|x^{(e)} - x^{(r_0)}\|, \quad \|x^{(e)} - x^{(i)}\| = O\big((1 - \|\,\mathbb{E}(X)\|)^2\big) \tag{7}$$

with the *concentration parameter* $1 - \|\,\mathbb{E}(X)\|$.

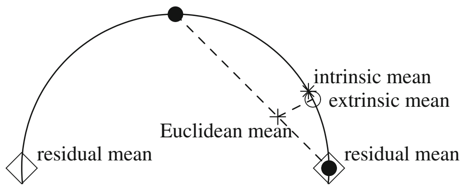

Fig. 2 Means on a circle of a distribution taking the *upper dotted value* with probability $1/3$ and the *lower right dotted value* with probability $2/3$. The latter happens to be one of the two residual means

Table 2 Mutual shape distances between intrinsic mean $\mu^{(i)}$, Ziezold mean $\mu^{(z)}$ and full Procrustes mean $\mu^{(p)}$ for various data sets

Data set	$d_{\Sigma_m^k}(\mu^{(i)},\mu^{(z)})$	$d_{\Sigma_m^k}(\mu^{(p)},\mu^{(z)})$	$d_{\Sigma_m^k}(\mu^{(p)},\mu^{(i)})$	$(1-\|\mathbb{E}(Y)\|)^2$
Poplar leaves	6.05e−05	1.83e−04	2.44e−04	5.24e−05
Digits ‘3’	0.00154	0.00452	0.00605	0.00155
Macaque skulls	1.96e−05	5.89e−05	7.85e−05	7.59e−06
Iron age brooches	0.000578	0.001713	0.002291	0.000217

Last column: the concentration parameter from (7), cf. also Corollary 4

Corollary 3 *Under condition (7), if all three means are unique, then the great circular segment between the residual mean $x^{(r_0)}$ closer to the extrinsic mean $x^{(e)}$ and the intrinsic mean $x^{(i)}$ is divided by the extrinsic mean in approximation by the ratio 1:3:*

$$x^{(r_0)} = \frac{\|\mathbb{E}(X)\|}{\lambda^{(r)}}\left(x^{(e)} - \frac{\mathbb{E}\big(\langle X - x^{(e)}, X\rangle X\big)}{\|\mathbb{E}(X)\|} + O\big((1-\|\mathbb{E}(X)\|)^2\big)\right)$$

$$x^{(i)} = \frac{\|\mathbb{E}(X)\|}{\lambda^{(i)}}\left(x^{(e)} + \frac{1}{3}\,\frac{\mathbb{E}\big(\langle X - x^{(e)}, X\rangle X\big)}{\|\mathbb{E}(X)\|} + O\big((1-\|\mathbb{E}(X)\|)^2\big)\right)$$

with $\lambda^{(i)}$ and $\lambda^{(r)}$ from Theorem 12.

Proof For any $x, p \in S^{D-1}$ decompose $p - x = p - \langle x, p\rangle x - z(x,p)x$ with $z(x,p) = 1 - \langle x, p\rangle$, the length of the part of p normal to the tangent space at x. Note that $\mathbb{E}\left(z(x^{(e)}, X\right) = 1 - \|\mathbb{E}(X)\|$. Now, under condition (7), verify the first assertion using Theorem 12:

$$x^{(r0)} = \frac{1}{\lambda^{(r)}}\left(\mathbb{E}(X) - \mathbb{E}\left(z(x^{(r_0)}, X)X\right)\right).$$

On the other hand since

$$\frac{\arccos(1-z)}{\sqrt{1-(1-z)^2}} = 1 + \frac{1}{3}\,z + \frac{2}{15}\,z^2 + \ldots$$

we obtain with the same argument the second assertion

$$\begin{aligned}
x^{(i)} &= \frac{1}{\lambda^{(i)}}\,\mathbb{E}\left(\frac{\arccos\langle X, x^{(i)}\rangle}{\sqrt{1-\langle X, x^{(i)}\rangle^2}}\,X\right)\\
&= \frac{1}{\lambda^{(i)}}\left(\mathbb{E}(X) + \frac{1}{3}\,\mathbb{E}\left(z(x^{(i)}, X)\,X\right) + \frac{2}{15}\,\mathbb{E}\left(z(x^{(i)}, X)^2\,X\right) + \ldots\right)\\
&= \frac{\|\mathbb{E}(X)\|}{\lambda^{(i)}}\left(x^{(e)} + \frac{1}{3\|\mathbb{E}(X)\|}\,\mathbb{E}\left(z(x^{(e)}, X)\,X\right) + O\big((1-\|\mathbb{E}(X)\|)^2\big)\right).
\end{aligned}$$

□

Remark 9 The tangent vector defining the great circle approximately connecting the three means is obtained from correcting with the expected normal component of any of the means. As numerical experiments show, this great circle is different from the first *principal component geodesic* as defined in Huckemann and Ziezold (2006).

Recall the following connection between top and quotient space means (cf. Theorem 4).

Remark 10 Let $p \in S_m^k$ such that a random shape X on Σ_m^k is supported by $(\Sigma_m^k)^{([p])} \cup A$ with $A \subset \Sigma_m^k$ at most countable. By Lemma 1 and Theorem 4, $(\Sigma_m^k)^{([p])} \cup A$ admits a horizontal measurable lift $L \subset S_m^k$ in optimal position to $p \in S_m^k$. Define the random variable Y on $L \subset S_m^k$ by $\pi \circ Y = X$. Then we have that

if $[p]$ is an intrinsic mean of X then p is an intrinsic mean of Y,
if $[p]$ is a full Procrustean mean of X then p is a residual mean of Y,
if $[p]$ is a Ziezold mean of X then p is an extrinsic mean of Y.

In consequence, Corollary 3 extends at once to Kendall's shape spaces. *Generalized geodesics* referred to below are an extension of the concept of geodesics to non-manifold shape spaces (cf. Huckemann et al. 2010b).

Corollary 4 *Suppose that a random shape X on Σ_m^k with unique intrinsic mean $\mu^{(i)}$, unique Ziezold mean $\mu^{(z)}$ and unique Procrustean mean $\mu^{(p)}$ is supported by $R = (\Sigma_m^k)^{(\mu^{(i)})} \cap (\Sigma_m^k)^{(\mu^{(z)})} \cap (\Sigma_m^k)^{(\mu^{(p)})}$. If the means are sufficiently close to each other in the sense of*

$$d_{\Sigma_m^k}(\mu^{(z)} - \mu^{(p)}),\ \ d_{\Sigma_m^k}(\mu^{(z)} - \mu^{(i)}) \ = \ O\big((1 - \|\mathbb{E}(Y)\|)^2\big)$$

with the random pre-shape Y on a horizontal lift L of R defined by $X = \pi \circ Y$, then the generalized geodesic segment between $\mu^{(i)}$ and $\mu^{(p)}$ is approximately divided by $\mu^{(z)}$ by the ratio 1:3 *with an error of order $O\big((1 - \|\mathbb{E}(Y)\|)^2\big)$.*

7 Examples: exemplary datasets and simulations

All of the results of this section are based on datasets and simulations, i.e., all means considered are sample means. An R-package for the computation of all means including the poplar leaves data can be found under Huckemann (2010).

7.1 The 1:3 property

In the first example we illustrate Corollary 4 on the basis of four classical data sets:

Poplar leaves: contains 104 quadrangular planar shapes extracted from poplar leaves in a joint collaboration with Institute for Forest Biometry and Informatics at the University of Göttingen (cf. Huckemann 2010; Huckemann et al. 2010a).

Digits '3': contains 30 planar shapes with 13 landmarks each, extracted from handwritten digits '3' (cf. Dryden and Mardia 1998, p. 318).

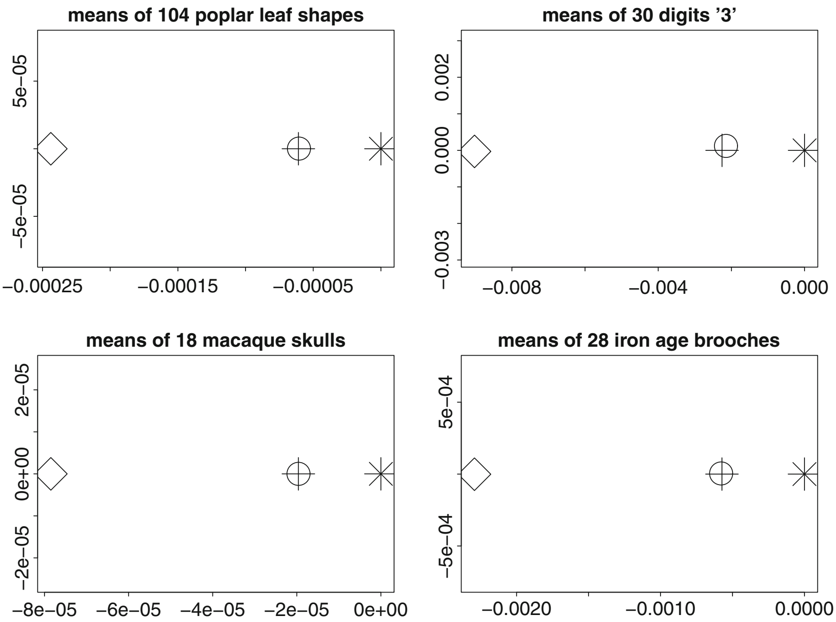

Fig. 3 Depicting shape means for four typical data sets: intrinsic (*star*), Ziezold (*circle*) and full Procrustes (*diamond*) projected to the tangent space at the intrinsic mean. The *cross* divides the generalized geodesic segment joining the intrinsic with the full Procrustes mean by the ratio 1:3

Macaque skulls:	contains three-dimensional shapes with 7 landmarks each, of 18 macaque skulls (cf. Dryden and Mardia 1998, p. 16).
Iron age brooches:	contains 28 three-dimensional tetrahedral shapes of iron age brooches (cf. Small 1996, Section 3.5).

As clearly visible from Fig. 3 and Table 2, the approximation of Corollary 4 for two- and three-dimensional shapes is highly accurate for data of little dispersion (the macaque skull data) and still fairly accurate for highly dispersed data (the digits '3' data).

7.2 "Partial Blindness" of full Procrustes and Schoenberg means

In the second example we illustrate an effect of "blindness to data" of full Procrustes means and Schoenberg means. The former blindness is due to the affinity of the Procrustes mean to the mode in conjunction with curvature, the latter is due to non-isometry of the Schoenberg embedding. While the former effect occurs only for some highly dispersed data when the analog of condition (7) is violated, the latter effect is local in nature and may occur for concentrated data as well.

Reenacting the situation of Sect. 4.4, cf. also Example 1 and Fig. 2, the shapes of the triangles q_1 and q_2 in Fig. 4 are almost maximally remote. Since the mode q_1 is assumed twice and q_2 only once, the full Procrustes mean is nearly blind to q_2.

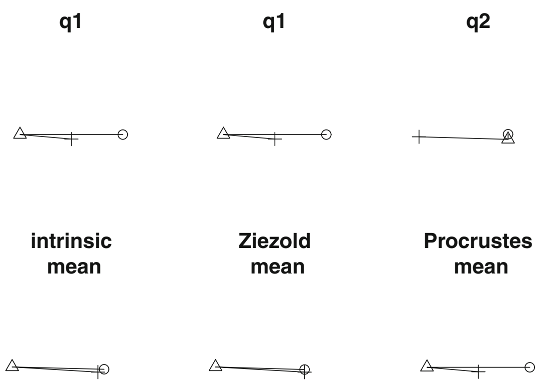

Fig. 4 A data set of three planar triangles (*top row*) with its corresponding intrinsic mean (*bottom left*), Ziezold mean (*bottom center*) and full Procrustes mean (*bottom right*)

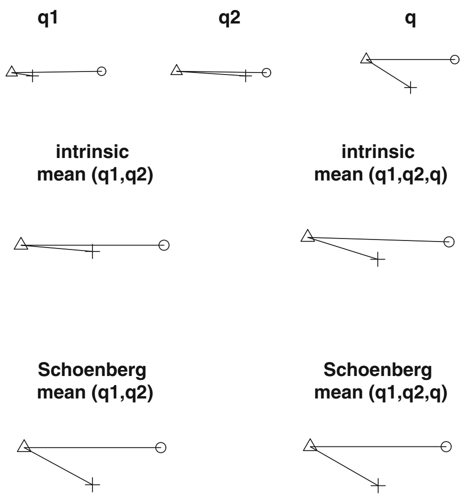

Fig. 5 Planar triangles $q_1 = x\cos\beta - w_2\sin\beta$, $q_2 = x\cos\beta + w_2\sin\beta$ and $q = x\cos\beta + w_1\sin\beta$ with x, w_1, w_2 from Remark 6, $\phi = 0.05$, $\beta = 0.3$ (*top row*). Intrinsic means (*middle row*) of sample (q_1, q_2) (*left*) and (q_1, q_2, q) (*right*). Schoenberg means (*bottom row*) of sample (q_1, q_2) (*left*) and (q_1, q_2, q) (*right*)

Even though Schoenberg means have been introduced to tackle 3D shapes, the effect of "blindness" can be well illustrated already for 2D. To this end consider $x = x(\phi)$, $w_1 = w_1(\phi)$ and $w_2 = w_2(\phi)$ as introduced in Remark 6. Along the horizontal geodesic through x with initial velocity w_2 we pick two points $q_1 = x\cos\beta + w_2\sin\beta$ and $q_2 = x\cos\beta - w_2\sin\beta$. On the orthogonal horizontal geodesic through x with initial velocity w_1 pick $q = x\cos\beta' + w_1\sin\beta'$. Recall from Remark 6, that along that geodesic the derivative of the Schoenberg embedding can be made arbitrarily small for ϕ near 0. Indeed, Fig. 5 illustrates that in contrast to the intrinsic mean, the Schoenberg mean is "blind" to the strong collinearity of q_1 and q_2.

7.3 Discrimination power

In the ultimate example we illustrate the consequences of the choice of tangent space coordinates and the effect of the tendency of the Schoenberg mean to increase

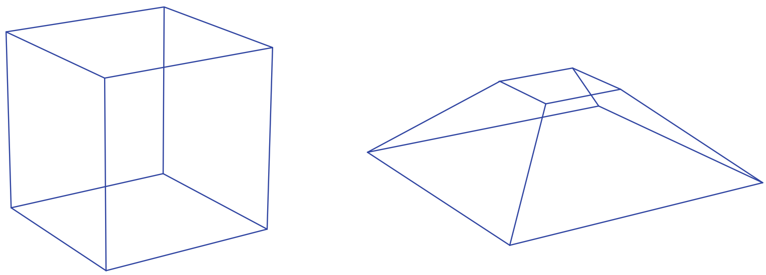

Fig. 6 Cube (*left*) and pyramid of varying height ϵ (*right*) for classification

Table 3 Percentage of correct classifications within 1,000 simulations each of 10 unit-cubes and 10 pyramids determined by ϵ (which gives the height), where each landmark is independently corrupted by Gaussian noise of variance $\sigma^2 = 0.2$ via a Hotelling T^2 test for equality of means to the significance level 0.05

ϵ	Intrinsic mean with intrinsic tangent space coordinates (%)	Intrinsic mean with residual tangent space coordinates (%)	Ziezold mean (%)	Schoenberg mean (%)
0.0	70	74	74	64
0.2	56	58	57	51
0.3	41	42	42	42

dimension by a classification simulation. To this end we apply a Hotelling T^2-test to discriminate the shapes of 10 noisy samples of regular unit cubes from the shapes of 10 noisy samples of pyramids with top section chopped off, each with 8 landmarks, given by the following configuration matrix

$$\begin{pmatrix} 0 & 1 & \frac{1+\epsilon}{2} & \frac{1-\epsilon}{2} & 0 & 1 & \frac{1+\epsilon}{2} & \frac{1-\epsilon}{2} \\ 0 & 0 & \frac{1-\epsilon}{2} & \frac{1-\epsilon}{2} & 1 & 1 & \frac{1+\epsilon}{2} & \frac{1+\epsilon}{2} \\ 0 & 0 & \epsilon & \epsilon & 0 & 0 & \epsilon & \epsilon \end{pmatrix}$$

(cf. Fig. 6) determined by $\epsilon > 0$. In the simulation, independent Gaussian noise is added to each landmark measurement. Table 3 gives the percentages of correct classifications. As visible from Table 3, discriminating flattened pyramids ($\epsilon = 0$) from cubes ($\epsilon = 1$) is achieved much better by employing intrinsic or Ziezold means rather than Schoenberg means. This finding is in concord with Theorem 7: samples of size 10 of two-dimensional configurations yield Euclidean means a.s. in $\mathcal{P}^7$ which are projected to $\mathcal{P}^3$ to obtain Schoenberg means in Σ_3^8. In consequence, Schoenberg means of noisy nearly two-dimensional pyramids are essentially three dimensional. With increased height of the pyramid ($\epsilon > 0$, i.e. for more pronounced third dimension and increased proximity to the unit cube) this effect waynes and all means perform equally well (or bad). Moreover in any case, intrinsic means with intrinsic tangent space coordinates qualify less for shape discrimination than intrinsic means with residual tangent space coordinates, cf. Remark 7. The latter (intrinsic means with residual tangent space coordinates) are better or equally well behaved as Ziezold means (which naturally use residual tangent space coordinates).

Table 4 Average time (s) for the computation of means in Σ_3^8 of sample size 20 on a PC with a 800 MHZ CPU based on 1, 000 repetitions

Intrinsic mean	Ziezold mean	Schoenberg mean
0.24	0.18	0.04

In conclusion, we record the time for the computations of means in Table 4. While Ziezold means compute in approximately $3/4$ of the computational time for intrinsic means, Schoenberg means are obtained approximately 6 times faster.

8 Discussion

By establishing stability results for intrinsic and Ziezold means on the manifold part of a shape space, a gap in asymptotic theory for general non-manifold shape spaces could be closed, now allowing for multi-sample tests of equality of intrinsic means and Ziezold means. A similar stability assertion in general is false for Procrustean means for low concentration. There is reason to believe, however, that it would be true for higher concentration. Note that the argument applied to intrinsic and Ziezold means fails for Procrustean means, since in contrast to the equations in Theorem 4 the sum of Procrustes residuals is in general non-zero. Loosely speaking, the findings on dimensionality condense to

- *Procrustean means may decrease dimension by 2 or more,*
- *intrinsic and Ziezold means decrease dimension at most by 1, in particular, they preserve regularity,*
- *Schoenberg means tend to increase up to the maximal dimension possible.*

Owing to the proximity of Ziezold and intrinsic means on Kendall's shape spaces in most practical applications, taking into consideration that the former are computationally easier accessible (optimally positioning and Euclidean averaging in every iteration step) than intrinsic means (optimally positioning and weighted averaging in every iteration step), Ziezold means can be preferred over intrinsic means. They may be even more preferred over intrinsic means, since Ziezold means naturally come with residual tangent space coordinates which may allow in case of intrinsic means for a higher finite power of tests than intrinsic tangent space coordinates.

Computationally much faster (not relying on iteration at all) are Schoenberg means which are available for Kendall's reflection shape spaces. As a drawback, however, Schoenberg means seem less sensitive for dimensionality of configurations considered than intrinsic or Ziezold means. In particular for problems involving small sample sizes, n and a large number of parameters p as currently of high interest in statistical applications, involving (nearly) degenerate data, Ziezold means may also be preferred over Schoenberg means due to higher power of tests.

Finally, note that Ziezold means may be defined for the shape spaces of planar curves introduced by Zahn and Roskies (1972), which are currently of interest, e.g. Klassen et al. (2004) or Schmidt et al. (2006). Employing Ziezold means there, a computational advantage greater than found here can be expected since the computation of iterates of intrinsic means involves computations of geodesics which themselves can only be found iteratively.

Appendix A: Proofs

Proof of Theorem 1. The assertion follows from Proposition 2, Remark 10 and the fact that $R\Sigma_j^k \subset \Sigma_m^k$ contains all shapes in Σ_m^k of configurations of dimension up to j, $1 \le j < m < k$. □

Proof of Theorem 2. Lemma 1 teaches that for Kendall's shape spaces, all quotient cut loci are void. Since for Ziezold means, Remark 2 provides invariant optimal positioning and Remark 3 provides the validity of (5), Corollary 1 applied to $R\Sigma_j^k$ as well as to Σ_m^k states that intrinsic and Ziezold means are also assumed on the manifold parts of $R\Sigma_j^k$ and Σ_m^k, respectively. In conjunction with Theorem 1, this gives the assertion. □

Lemma 2 *Let $U \subset M$ be a tubular neighborhood about $p \in M$ that admits a slice via $\exp_p D \times_{I_p} G \cong U$ in optimal position to p. Then, there is a measurable horizontal lift $L \subset \exp_p D$ of $\pi(U)$ in optimal position to p.*

Proof If p is regular, then $L = \exp_p D$ has the desired properties. Now assume that p is not of maximal orbit type. W.l.o.g. assume that D contains the closed ball B of radius $r > 0$ with bounding sphere $S = \partial B$ and that there are $p^1, \dots, p^J \in \exp_p(S)$ having the distinct orbit types occurring in S. S^j denotes all points on S of orbit type (G/I_{p^j}), $j = 1, \dots, J$, respectively. Observe that each S^j is a manifold on which I_p acts isometrically. Hence for every $1 \le j \le J$, there is a finite ($K_j < \infty$) or countable ($K_j = \infty$) sequence of tubular neighborhoods $U_k^j \subset S^j$ covering S^j, admitting trivial slices via

$$\exp^{S^j}_{p_k^j} D_k^j \times I_p/I_{p_k^j} \cong U_k^j, \quad p_k^j \in U_k^j, \quad 1 \le k \le K_j.$$

Here, $\exp^{S^j}_{p_k^j}$ denotes the Riemannian exponential of S^j. Defining a disjoint sequence

$$\widetilde{U}_1^j := U_1^j, \quad \widetilde{U}_{k+1}^j := U_{k+1}^j \setminus \widetilde{U}_k^j \text{ for } 1 \le k \le K_j - 1$$

exhausting S^j we obtain a corresponding sequence of disjoint measurable sets $\exp^{S^j}_{p_k^j} \widetilde{D}_k^j$ with

$$\exp^{S^j}_{p_k^j} \widetilde{D}_k^j \times I_p/I_{p_k^j} \cong \widetilde{U}_k^j, \quad 1 \le k \le K_j.$$

Hence, setting

$$L_k^j := \exp_{p_k^j}^{S^j} \widetilde{D}_k^j \text{ and } L^j := \bigcup_{k=1}^{K_j} L_k^j$$

observe that every $p' \in S^j$ has a unique lift in L^j which is contained in a unique L_k^j. This lift is by construction (all L_k^j are in $\exp_p D$) in optimal position to p. Moreover, if $p' \in L_k^j$ and $gp' \in L_{k'}^j$ for some $g \in G$ with $1 \leq k', k \leq K_j$ we have by the disjoint construction of $\widetilde{U}_k^j$ and $\widetilde{U}_{k'}^j$ that $k = k'$, hence the isotropy groups of gp' and p' agree, yielding $gp' = p'$. In consequence, L^j is a measurable horizontal lift of S^j in optimal position to p. Since every horizontal geodesic segment $t \mapsto \exp_p(tv)$, $v \in H_pM$ contained in $\exp_p D$ features a constant isotropy group, except possibly for the initial point we obtain with the definition of

$$L := \cup_{j=1}^J M^j \text{ with } M^j := \{\exp_p(tv) \in \exp_p(D) : v \in \exp_p^{-1}(L^j), t \geq 0\}$$

a measurable horizontal lift of $\pi(U)$ in optimal position to p. □

Proof of Theorem 3. Since M is connected, any two points p, p' can be brought into optimal position p, gp' and a closed minimizing horizontal geodesic segment $\gamma_{gp'}$ between p, gp' can be found. If $[p'] \in Q^{([p])}$ then also $\gamma_{gp'} \subset Q^{([p])}$. In consequence, there are tubular neighborhoods U_p of p and $U_{p'}$ of $\gamma_{gp'}$ admitting slices in optimal position to p, which by Lemma 2, have horizontal lifts L_p and $L_{p'}$ in optimal position to p. Since M is a manifold, there is a sequence $[p_0], \ldots \in Q^{([p])}$, $g_j \in G$, $p_j \in M$ such that $p_0 = p$ and that each $g_j p_j$ is in optimal position to p ($j \in J,\ J \subset \mathbb{N}$) and such that

$$Q^{([p])} \subset \bigcup_{j \in J \cup \{0\}} \pi(U_{p_j})$$

with measurable horizontal lifts L_{p_j} of $\pi(U_{p_j})$. Defining $L'_{p_0} := L_{p_0}$ and recursively $L'_{p_{j+1}} := L_{p_{j+1}} \setminus L'_{p_j}$ for $j = 1, \ldots$ a measurable horizontal lift $L' := \cup_{j=0}^\infty L'_{p_j}$ of $Q^{([p])}$ in optimal position to p is obtained. Finally, suppose that p_j is in optimal position to p for $p_j \in [p_j] \in A$ and set $L''_0 := L', L''_j := L''_{j-1} \cup \{p_j\}$ if $[p_j] \cap L'_j = \emptyset$ and $L''_j = L''_{j-1}$ ($j \geq 1$) otherwise to obtain the desired measurable horizontal lift $L'' := \cup_{[p_j] \in A} L''_j$ in optimal position to p. □

Proof of Theorem 5. In case of intrinsic means, with the hypotheses and notations of the above proof of Theorem 3, suppose that L'' is a measurable horizontal lift of $Q^{([p])} \cup A$ in optimal position to an intrinsic mean $p \in M$ of the random element Y on M defined as in Theorem 4 with $[p] \in E^{(d_Q)}(X)$. For notational simplicity we assume that $Q^{([p])} = \pi(U)$ with a single tubular neighborhood U of p admitting a slice.

Then, additionally using the notation of the above proof of Lemma 2, if the assertion of the Theorem were false, w.l.o.g. there would be $g \in I_p$, $1 \leq j \leq J$, $p_j \in S^j$ with $gp_j \neq p_j$ and $\mathbb{P}\{Y \in M^j\} > 0$. In particular, in the proof Lemma 2, we may choose a sufficiently small U_k^j around p_j such that in consequence of (4)

$$\int_{M_k^j(\epsilon)} \left(\exp_p^{-1} Y - \exp_p^{-1}(gY)\right) d\,\mathbb{P}_Y \neq 0 \tag{8}$$

with some $\epsilon, r > 0$, $M_k^j(\epsilon) := \{\exp_p(tv) \in \exp_p(D) : v \in \exp_p^{-1}(L_k^j), |t - r| < \epsilon\}$ and L_k^j obtained from U_k^j as in the proof of Lemma 2. Suppose that $L \subset L''$ is obtained as in the proof of Lemma 2 by using L_k^j and suppose that $\widetilde{L''}$ is obtained from L'' by replacing the $M_k^j(\epsilon)$ part of M_k^j with $\{\exp_p(tv) \in \exp_p(D) : v \in \exp_p^{-1}(gL_k^j), |t - r| < \epsilon\}$. Then $\widetilde{L''}$ is also a measurable horizontal lift in optimal position to p. Since we assume that $[p]$ is an intrinsic mean of X, assertion (i) of Theorem 4 teaches that p is also an intrinsic mean of lift $\widetilde{Y}$ of X to $\widetilde{L''}$, i.e.

$$\begin{aligned} 0 &= \int_{L''} \exp_p^{-1} Y\, d\,\mathbb{P}_Y - \int_{\widetilde{L''}} \exp_p^{-1} \widetilde{Y}\, d\,\mathbb{P}'_Y \\ &= \int_{M_k^j(\epsilon)} \left(\exp_p^{-1} Y - \exp_p^{-1}(gY)\right) d\,\mathbb{P}_Y\,. \end{aligned}$$

This is a contradiction to (8) yielding the validity of the theorem for intrinsic means.

The assertion in case of Ziezold means is similarly obtained. Use the same horizontal lifts L'' and $\widetilde{L''}$ from above, replace $\exp_p^{-1} Y$, $\exp_p^{-1} \widetilde{Y}$ and $\exp_p^{-1}(gY)$ by $df_{\mathrm{ext}}^{Y}(p)$, $df_{\mathrm{ext}}^{\widetilde{Y}}(p)$ and $df_{\mathrm{ext}}^{gY}(p)$, respectively, use the hypothesis (5) to obtain the analog of (8) and finally obtain the contradiction arguing with assertion (ii) of Theorem 4. □

Acknowledgments The author would like to thank Alexander Lytchak for helpful advice on differential geometric issues. Also, the author gratefully acknowledges support by DFG Grant HU 1575/2-1.

References

Afsari, B. (2011). Riemannian L^p center of mass: existence, uniqueness, and convexity. *Proceedings of the American Mathematical Society, 139*, 655–773.

Anderson, T. (2003). *An introduction to multivariate statistical analysis* (3rd ed.). New York: Wiley.

Bandulasiri, A., Patrangenaru, V. (2005). Algorithms for nonparametric inference on shape manifolds. *Proceedings of JSM 2005 Minneapolis, MN*, 1617–1622.

Bhattacharya, A. (2008). Statistical analysis on manifolds: A nonparametric approach for inference on shape spaces. *Sankhya, Series A, 70*(2), 223–266.

Bhattacharya, R. N., Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds I. *The Annals of Statistics, 31*(1), 1–29.

Bhattacharya, R. N., Patrangenaru, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds II. *The Annals of Statistics, 33*(3), 1225–1259.

Billera, L., Holmes, S., Vogtmann, K. (2001). Geometry of the space of phylogenetic trees. *Advances in Applied Mathematics, 27*(4), 733–767.

Bredon, G. E. (1972). Introduction to compact transformation groups. In *Pure and applied mathematics* (Vol. 46). New York: Academic Press.

Choquet, G. (1954). Theory of capacities. *Annales de l'Institut de Fourier, 5*, 131–295.
Dryden, I. L., Mardia, K. V. (1998). *Statistical shape analysis*. Chichester: Wiley.
Dryden, I. L., Kume, A., Le., H., Wood, A. T. A. (2008). A multidimensional scaling approach to shape analysis (to appear).
Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *10*(4), 215–310.
Gower, J. C. (1975). Generalized Procrustes analysis. *Psychometrika, 40,* 33–51.
Hendriks, H., Landsman, Z. (1996). Asymptotic behaviour of sample mean location for manifolds. *Statistics and Probability Letters, 26*, 169–178.
Hendriks, H., 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., Ruymgaart, F. (1996). Asymptotic behaviour of sample mean direction for spheres. *Journal of Multivariate Analysis, 59*, 141–152.
Hotz, T., Huckemann, S., Le, H., Marron, J. S., Mattingly, J. C., Miller, E., Nolen, J., Owen, M., Patrangenaru, V., Skwerer, S. (2012). Sticky central limit theorems on open books. arXiv.org, 1202.4267 [math.PR] [math.MG] [math.ST].
Huckemann, S. (2010). R-package for intrinsic statistical analysis of shapes. http://www.mathematik.uni-kassel.de/~huckeman/software/ishapes_1.0.tar.gz.
Huckemann, S. (2011). Inference on 3D Procrustes means: Tree boles growth, rank-deficient diffusion tensors and perturbation models. *Scandinavian Journal of Statistics, 38*(3), 424–446.
Huckemann, S., Ziezold, H. (2006). Principal component analysis for Riemannian manifolds with an application to triangular shape spaces. *Advances of Applied Probability (SGSA), 38*(2), 299–319.
Huckemann, S., Hotz, T., Munk, A. (2010a). Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(4), 593–603.
Huckemann, S., Hotz, T., Munk, A. (2010b). Intrinsic shape analysis: Geodesic principal component analysis for Riemannian manifolds modulo Lie group actions (with discussion). *Statistica Sinica, 20*(1), 1–100.
Jupp, P. E. (1988). Residuals for directional data. *Journal of Applied Statistics, 15*(2), 137–147.
Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 509–541.
Kendall, D. (1974). Foundations of a theory of random sets. In *Stochastic geometry, tribute memory Rollo Davidson* (pp. 322–376). New York: Wiley.
Kendall, D. G., Barden, D., Carne, T. K., Le, H. (1999). *Shape and shape theory*. Chichester: Wiley.
Kendall, W. S. (1990). Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence. *Proceedings of the London Mathematical Society, 61*, 371–406.
Kent, J., Hotz, T., Huckemann, S., Miller, E. (2011). The topology and geometry of projective shape spaces (in preparation).
Klassen, E., Srivastava, A., Mio, W., Joshi, S. (2004, March). Analysis on planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 26*(3), 372–383.
Kobayashi, S., Nomizu, K. (1963). *Foundations of differential geometry* (Vol. I). Chichester: Wiley.
Kobayashi, S., Nomizu, K. (1969). *Foundations of differential geometry* (Vol. II). Chichester: Wiley.
Krim, H., Yezzi, A. J. J. E. (2006). Statistics and analysis of shapes. In *Modeling and Simulation in Science, Engineering and Technology*. Boston: Birkhäuser.
Le, H. (2001). Locating Fréchet means with an application to shape spaces. *Advances of Applied Probability (SGSA), 33*(2), 324–338.
Le, H. (2004). Estimation of Riemannian barycenters. *LMS Journal of Computation and Mathematics, 7*, 193–200.
Lehmann, E. L. (1997). Testing statistical hypotheses. In *Springer texts in statistics*. New York: Springer.
Mardia, K., Patrangenaru,. V. (2001). On affine and projective shape data analysis. In K. V. Mardia, R. G. Aykroyd (Eds.), *Proceedings of the 20th LASR Workshop on functional and spatial data analysis* (pp. 39–45).
Mardia, K., Patrangenaru, V. (2005). Directions and projective shapes. *The Annals of Statistics, 33*, 1666–1699.
Matheron, G. (1975). Random sets and integral geometry. In *Wiley series in probability and mathematical statistics*. New York: Wiley.

Nash, J. (1956). The imbedding problem for Riemannian manifolds. *The Annals of Mathematics, 63*, 20–63.

Palais, R. S. (1961). On the existence of slices for actions of non-compact Lie groups. *The Annals of Mathematics 2nd Series, 73*(2), 295–323.

Schmidt, F. R., Clausen, M., Cremers, D. (2006). Shape matching by variational computation of geodesics on a manifold. In *Pattern recognition (Proceedings DAGM), LNCS* (Vol. 4174, pp. 142–151). Berlin: Springer.

Small, C. G. (1996). *The statistical theory of shape*. New York: Springer.

Zahn, C., Roskies, R. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions on Computers, C 21*, 269–281.

Ziezold, H. (1977). Expected figures and a strong law of large numbers for random elements in quasi-metric spaces. In *Transaction of the 7th Prague conference on information theory, statistical decision function and random processes, A*, pp. 591–602.

Ziezold, H. (1994). Mean figures and mean shapes applied to biological figure and shape distributions in the plane. *Biometrical Journal, 36*, 491–510.