Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Set-valued means of random particles

D. Stoyan and I.S. Molchanov

Department of Operations Reasearch, Statistics, and System Theory

BS-R9511 1995

Report BS-R9511
ISSN 0924-0659

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Set-Valued Means of Random Particles

Dietrich Stoyan
TU Bergakademie Freiberg, Institut für Stochastik
D-09596 Freiberg, Germany
Stoyan@mathe.tu-freiberg.d400.de

I.S. Molchanov
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
ilia@cwi.nl

Abstract

Planar images of powder particles or sand grains can be interpreted as "figures", i. e. equivalence classes of directly congruent compact sets. The paper introduces a concept of set-valued means and real-valued variances for samples of such figures. In obtaining these results, the images are registered to have similar locations and orientations. The method is applied to find a mean figure of a sample of polygonal particles.

AMS Subject Classification (1991): 60D05, 62D05, 68U10.

Keywords & Phrases: estimation, expectation, figure, orbit space, particle, random set, shape, support function, variance.

Notes: This research was partially supported by the Alexander von Humboldt-Stiftung, Bonn, Germany. The first author was supported also by the Netherlands Organization for Scientific Research (NWO). This paper has been submitted for publication.

1. Introduction

Particle characterization is an important problem of particle technology and sedimentology. It includes the statistical analysis of samples of planar objects such as projections of powder particles or sand grains, which can be considered as planar compact sets. The traditional way to deal with such data is to use motion-invariant shape and size characteristics, which typically are based on area, perimeter and diameter etc., see for example [14] and [21]. These values are analyzed by the usual methods of multivariate statistics.

This paper suggests an approach from the area of set theory and produces a compact set as the mean of a sample of particles. The problem is difficult since orientation and location of the particles are arbitrary; directly congruent particles are considered as identical. Consequently, the existing theories of mean values of random compact sets are not directly applicable. Our approach is inspired by studies of shapes and landmark configurations, see [4], [8], [17], [25] and [26]. Landmarks are characteristic points of planar objects, such as the tips of the nose and the chin, if human profiles are studied. It is important that all objects in a sample have the same number of landmarks and that landmarks of the same order have the same meaning. However, for the study of particles such landmarks are not natural. Perhaps they could be points of extremal curvature or other interesting points on the boundary, but for a useful application of the landmark method the number of landmarks per object has to be constant, and this may lead to difficulties or unnatural restrictions.

For the practically important problem of determining an empirical mean of n figures the following general idea seems to be natural:

> Give the figures particular locations and orientations such that they are in a certain sense "close together"; then consider the new sample as a sample of sets and, finally, determine a set-theoretic mean.

This idea appears in [10]. L. Galway dealt with so-called radial averages of star-shaped sets and carried out a statistical analysis of sand grains in the following manner:

> "The first problem is to locate the origin within the kernel of each grain .... Each grain has been smoothed to have a non-null and full-dimensional kernel, so the problem reduces to locating the origin within the kernel. The choice was made to centre each grain at the centroid (= centre of gravity) of the kernel ..."
>
> "Orientation is more difficult ... random orientation would change the expected set by introducing a new source of randomness in addition to the distribution of shapes. However, the only clues to orientation are the grain profiles themselves. A two stage procedure was followed: after centring all grains, the samples were broken into subsamples of about fifty. Each subsample was rotated to an angle where the rotated profile had the smallest Hausdorff distance to the first grain in the subsample. Each subsample was radially averaged, the radial averages were aligned (by rotating each to have the minimum Hausdorff distance to the radial average of the first subsample) and then the subsample averages were radially averaged to produce a sample radial average. This sample average was then used to realign the whole sample once more and then a final sample radial average was computed."

This paper tries to justify this approach. The particles are considered as equivalence classes of elements of a Hilbert space equipped with a group of transformations. Elementary methods of Hilbert space theory lead to a characterization of a suitably defined mean which is the basis of the statistical procedure. A numerical example with polygonal particles demonstrates its application.

2. Random Compact Sets and Their Means

This section presents the concepts of mean and variance which we shall use for the shifted and rotated samples. Furthermore, we give three examples of particular practical interest. Let $\mathcal{K}'$ be the system of all non-empty compact sets in $\mathbb{R}^d$. A *random compact set* X is a random element with values in $\mathcal{K}'$ endowed with the Borel σ-algebra corresponding to the Hausdorff metric h given by

$$h(K_1, K_2) = \inf\{r > 0 : K_1 \subset K_2^r, K_2 \subset K_1^r\},$$

where $K^r = K \oplus b(o, r)$ denotes the closed r-neighbourhood of K, where $\oplus$ is the Minkowski addition and $b(o, r)$ is the ball of radius r centred at the origin o. For more details see [19] and [21].

The space $\mathcal{K}'$ of compact sets is not linear. This makes it difficult to define directly a mean of a random set. All existing definitions of means of random compact sets use random functions which characterize the sets. There a random compact set X corresponds to a random function $\xi_X(t)$, $t \in M$, where M is $\mathbb{R}^d$ or a subset of a lower-dimensional Euclidean space, and its mean in a certain space H of functions can be determined. For defining the mean we use the Fréchet approach ([9], [21, p. 112]). In this approach a mean of a random element η of a metric space H with metric ρ is an element a of H with

$$\mathbf{E}((\rho(\eta, a))^2) \to \min!$$

In general, η may have several means. The *variance* of η is the minimum value of $\mathbf{E}((\rho(\eta, \cdot))^2)$.

We consider here only the case where H is a Hilbert space. Then the Fréchet approach yields the usual mean. If, in particular, H is the Hilbert space $L^2(M)$, then this mean is equal to the expectation $\mathbf{E}(\xi_X(t))$, $t \in M$, which is again a function. If this non-random function corresponds to a deterministic set, then this set is naturally called the expectation of X. If this is not the case, then it is possible either to construct some sets from $\mathbf{E}(\xi_X(t))$ (e. g. sets given by contour lines) or to use this function itself as a mean.

Let us consider three examples.

Example. (Aumann expectation)
Let X be convex and let ξ_X be the support function of X defined as

$$\xi_X(t) = s_X(t) = \sup\{\langle t, u\rangle : u \in X\}, \quad t \in \mathbb{S}^{d-1},$$

where $\langle t, u\rangle$ is the scalar product in $\mathbb{R}^d$ and $\mathbb{S}^d$ is the unit sphere. Then $\mathbf{E}(s_X(\cdot))$ is again a support function, namely that of a deterministic *convex* set $\mathbf{E}X$, which is called the Aumann expectation of X, see e. g. [21]. Note that it is possible to define this expectation also in terms of selections or through the integral of a multivalued function, see [1], [22] and [21].

The corresponding variance is

$$\int_{\mathbb{S}^{d-1}} \mathbf{E}(\mathbf{E}(s_X(t)) - s_X(t))^2 \mathrm{d}t\,.$$

It is possible to interpret the Aumann mean as a Bochner integral in the Banach space of continuous functions on $\mathbb{S}^{d-1}$. The Fréchet approach for the space $L^2(\mathbb{S}^{d-1})$ yields the same result.

Example. (Vorob'ev expectation)
Let $\xi_X(x) = 1_X(x)$ be the indicator function of X. Then $\mathbf{E}(\xi_X(x)) = p_X(x) = \mathbf{P}(x \in X)$ is the *coverage function*. In general, $p_X(x)$ is not an indicator function. Nevertheless, it seems to be natural to use the function $p_X(x)$ as a "mean" of X. A set-theoretic mean is defined in [23] by

$$L_p = \{x \in \mathbb{R}^d : p_X(x) \geq p\}$$

for p which is determined by the inequality

$$\nu(L_q) \leq \mathbf{E}(\nu(X)) \leq \nu(L_p)\,, \quad \text{for all} \quad q > p$$

for the Lebesgue measure ν. The set $L_{1/2}$ has properties of a *median*, see [21] and [23].

This approach considers indicator functions as elements of $L^2(\mathbb{R}^d)$. It implies that singletons as well as sets of almost surely vanishing Lebesgue measure are considered as uninteresting, since the corresponding indicator random field $1_X(x)$ vanishes almost surely. The corresponding variance is

$$\int_{\mathbb{R}^d} \mathbf{E}(p_X(x) - \mathbf{1}_X(x))^2 \mathrm{d}x\,.$$

Example. (Radius-vector mean)
Let X be *shrinkable* with respect to the origin o, i. e. let $[0,1)X^{\mathrm{cl}} \subset X^{\mathrm{int}}$, where X^{cl} is the closed hull of X and X^{int} its interior. (A shrinkable set is also star-shaped.) Let r_X be the *radius-vector function* defined by

$$r_X(t) = \sup\{x : xt \in X, x \geq 0\}\,, \quad t \in \mathbb{S}^{d-1}\,.$$

The means $\mathbf{E}(r_X(t))$ define a function which can be considered as the radius-vector function of a deterministic shrinkable set, which is called the radius-vector mean of X, see [21].

In the planar case radius-vector functions are very popular in the engineering literature. There for shape description Fourier methods are applied, see e. g. [2].

Note that for these examples proper Euclidean motions (rotations for the last example) correspond to isometric transformations in the corresponding L^2 spaces.

3. Means in Orbit Spaces

Let H be a Hilbert space with norm $\|\cdot\|$. Furthermore, let G be a group acting on H from the left. For each $x \in H$ its *orbit* is the set

$$Gx = \{gx : g \in G\}\,.$$

By $x \sim y$ iff $y \in Gx$ an equivalence relation on H is defined. The corresponding factor space H/G is said to be the *orbit space*, see [5]. We assume that all $g \in G$ are *isometric*, i. e. $\|gx - gy\| = \|x - y\|$ for all $x,\ y \in H$ and all $g \in G$.

For each x and y in H define

$$\rho_G(x,y) = \inf_{g_1,g_2 \in G} \|g_1 x - g_2 y\| .$$

Clearly, it is

$$\rho_G(x,y) = \inf_{g \in G} \|gx - y\| .$$

This function induces a *metric* on the orbit space H/G, which is also denoted by ρ_G. In the following the factor space H/G is equipped with the Borel σ-algebra corresponding to this metric.

Statistical problems such as those discussed in the introduction can be formulated in the orbit space. If H is a space of functions describing compact sets and G is a group corresponding to proper Euclidean motions in $\mathbb{R}^d$, then the space H/G is the right space for statistics of particles. However, the orbit space is usually curved, which makes it difficult to apply routine statistical procedures to samples composed of its elements. The same problem appears in the statistical theory of shapes, see e. g. [6], [7] and [17].

Now we introduce some necessary notions related to samples in the orbit space.

A finite set of points $\boldsymbol{x} = \{x_1, \ldots, x_n\} \subset H$ is said to be a *configuration*. For each configuration define its *inertia* as

$$\mathcal{I}(\boldsymbol{x}) = \sum_{1 \le i < j \le n} \|x_i - x_j\|^2 .$$

Definition 1. The configuration $\boldsymbol{x}$ is said to be in *optimal position* (with respect to G) if and only if

$$\mathcal{I}(\boldsymbol{x}) = \inf_{g_1,\ldots,g_n \in G} \mathcal{I}(\boldsymbol{g}\boldsymbol{x}) ,$$

where $\boldsymbol{g}\boldsymbol{x} = \{g_1 x_1, \ldots, g_n x_n\}$.

A configuration $\boldsymbol{x}$ corresponds to a finite subset $\tilde{\boldsymbol{x}}$ of the orbit space H/G. This set consists of the equivalence classes generated by the elements of $\boldsymbol{x}$.

Now let us consider means of configurations.

Definition 2. The element $a \in H$ is a *mean* of the configuration $\boldsymbol{x}$ if

$$\mathcal{I}(\boldsymbol{x}, a) = \inf_{y \in H} \mathcal{I}(\boldsymbol{x}, y) .$$

Here $\mathcal{I}(\boldsymbol{x}, a)$ is the inertia of the configuration $\boldsymbol{x} \cup \{a\} = \{x_1, \ldots, x_n, a\}$.

This definition (in different terms) is taken from [24] and it is a particular case of the Fréchet expectation. If we use the fact that

$$\mathcal{I}(\boldsymbol{x}, y) = \mathcal{I}(\boldsymbol{x}) + \sum_{i=1}^{n} \|x_i - y\|^2 ,$$

then it is clear that $y = a$ minimizes

$$\sum_{i=1}^{n} \|x_i - y\|^2 = \mathcal{I}(\boldsymbol{x}, y) - \mathcal{I}(\boldsymbol{x})$$

for $y \in H$.

The following theorem can be proved quite easily using the properties of the norm in a Hilbert space.

Theorem 1. *The mean of the configuration $\boldsymbol{x}$ is unique and is given by*

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i .$$

Definition 3. The point $b \in H$ is said to be a *relative mean* of $\boldsymbol{x}$ (with respect to G) (notation: $b \in \overline{x}_G$) if

$$\inf_{g_1,\ldots,g_n \in G} \mathcal{I}(\boldsymbol{gx}, b) - \mathcal{I}(\boldsymbol{gx}) = \inf_{y \in H} \inf_{g_1,\ldots,g_n \in G} \mathcal{I}(\boldsymbol{gx}, y) - \mathcal{I}(\boldsymbol{gx}) . \tag{3.1}$$

An equivalent version of (3.1) is

$$\begin{aligned} \sum_{i=1}^{n} \rho_G(x_i, b)^2 &= \inf_{g_1,\ldots,g_n \in G} \sum_{i=1}^{n} \|g_i x_i - b\|^2 \\ &= \inf_{y \in H} \inf_{g_1,\ldots,g_n \in G} \sum_{i=1}^{n} \|g_i x_i - y\|^2 = \inf_{y \in H} \sum_{i=1}^{n} \rho_G(x_i, y)^2 . \end{aligned}$$

It is clear that $\overline{x}_G$ consists of all representatives of all equivalence classes belonging to $\text{mean}_G(\tilde{\boldsymbol{x}})$, where $\text{mean}_G(\tilde{\boldsymbol{x}})$ is the set of all Fréchet means of the sample $\tilde{\boldsymbol{x}}$ in H/G with respect to the metric ρ_G. Therefore, the relative mean corresponds to the mean in the orbit space H/G. The latter is of interest when analyzing samples in the orbit space.

Remark

(1) If b is a relative mean of $\boldsymbol{x}$, then also gb is a relative mean for any $g \in G$. Thus, a relative mean can be considered to be an element of the orbit space.

(2) For any $\boldsymbol{g}$, $\overline{(gx)}_G = \overline{x}_G$, where $\overline{(gx)}_G$ denotes the relative mean of $\boldsymbol{gx}$.

The following result is the basis of our method for obtaining means in orbit spaces.

Theorem 2. *The configuration $\boldsymbol{x}$ is in optimal position if and only if*

1. $\overline{x} \in \overline{x}_G$;
2. *for every i, the point x_i and $\overline{x}$ are in optimal position.*

The proof of Theorem 2 is given in the Appendix.

H. Karcher [15] introduced a mean in metric spaces through local infima as in Definition 2. Following this idea, it is possible to define *local* variants of all notions introduced above.

4. Characterizing Configurations in Optimal Position

Theorem 2 is a basis for determining means and variances of particles. It shows that it is necessary to bring the members of a sample of compact sets into optimal position. This is a large optimization problem. In the planar case $3(n-1)$ real numbers have to be determined: the $2(n-1)$ components of the shift vectors and $n-1$ angles.

Suppose that the orbit Gx is compact for each x. Then, for any x and $y \in H$, there exists an element $g_* \in G$ such that

$$\|g_* x - y\| = \inf_{g\in G} \|gx - y\| .$$

Set $g_*x = \phi(y, x)$. This notation is used in the following algorithm which may be used to transform a configuration $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ to its optimal position. It is in the spirit of Gower's generalized procrustes algorithm [12].

Algorithm 1.

1. Set $\boldsymbol{y} = \boldsymbol{x}$.
2. Compute $\overline{y}$.
3. Transform the configuration $\boldsymbol{y}$ to another configuration $\boldsymbol{y}'$ by replacing y_i with $y_i' = \phi(\overline{y}, y_i)$, $i = 1, \ldots, n$.
4. If $\boldsymbol{y}'$ is close to $\boldsymbol{y}$, then stop. Otherwise set $\boldsymbol{y} = \boldsymbol{y}'$ and go to 2.

Each step of this algorithm reduces the inertia of the configuration. Indeed,

$$\begin{aligned}\mathcal{I}(\boldsymbol{y}) &= n(\mathcal{I}(\boldsymbol{y},\overline{y}) - \mathcal{I}(\boldsymbol{y})) \;\geq\; n(\mathcal{I}(\boldsymbol{y}',\overline{y}) - \mathcal{I}(\boldsymbol{y}')) \\ &\geq n(\mathcal{I}(\boldsymbol{y}',\overline{y'}) - \mathcal{I}(\boldsymbol{y}')) \;=\; \mathcal{I}(\boldsymbol{y}') .\end{aligned}$$

The group structure can be exploited to simplify the search for the optimal position. For this simplification, suppose that each element $g \in G$ acts on H as the transformation

$$gx = x + l \tag{4.1}$$

with a corresponding element $l = l(g)$ of a linear subspace $\mathcal{L}$ of H. The map $g \mapsto l(g)$ defines an isomorphism between G and $\mathcal{L}$, where $\mathcal{L}$ is equipped with the addition operation. The projection of $x \in H$ on $\mathcal{L}$ is denoted by $\mathrm{pr}_{\mathcal{L}} x$.

Theorem 3. *If the group G is defined by (4.1), then the configuration $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ is in optimal position if and only if*

$$\mathrm{pr}_{\mathcal{L}} x_1 = \mathrm{pr}_{\mathcal{L}} x_2 = \cdots = \mathrm{pr}_{\mathcal{L}} x_n . \tag{4.2}$$

Proof. According to the Lemma in the Appendix,

$$\begin{aligned}&\frac{1}{n^2} \sum_{1 \leq i < j \leq n} \|x_i + l_i - x_j - l_j\|^2 \;= \\ &\quad = \frac{1}{n} \sum_{i=1}^{n} \left\| x_i + l_i - \frac{1}{n}\sum_{j=1}^{n} x_j - \frac{1}{n}\sum_{j=1}^{n} l_j \right\|^2 .\end{aligned}$$

Each summand in the right-hand side is minimal if

$$\mathrm{pr}_{\mathcal{L}} \left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j \right) = - \left(l_i - \frac{1}{n}\sum_{j=1}^{n} l_j \right), \quad i = 1, \ldots, n .$$

This system is solved by

$$l_i = b - \mathrm{pr}_{\mathcal{L}} x_i , \quad i = 1, \ldots, n$$

for any $b \in \mathcal{L}$. Notice that $l_i = 0,\ 1 \leq i \leq n$, if and only if $\boldsymbol{x}$ is in optimal position. Thus, $b = \mathrm{pr}_{\mathcal{L}} x_i$, for each i, which is equivalent to (4.2). □

Sometimes it is possible to decompose G in such a way that, for each $g \in G$ and $x \in H$,

$$gx = \tilde{g}x + l , \quad \tilde{g} \in \tilde{G} , \quad l = l(g) \in \mathcal{L} , \tag{4.3}$$

for a subgroup $\tilde{G}$ of G and a linear subspace $\mathcal{L}$ of H. For example, the group of all Euclidean motions in $H = \mathbb{R}^d$ admits such a representation, where $\tilde{G}$ is the group of rotations and $\mathcal{L} = H$. Let $\mathcal{L}^\perp$ be the orthogonal complement to $\mathcal{L}$.

Theorem 4. *If $\tilde{G}$ in (4.3) consists of linear operators, $\tilde{G}\mathcal{L} \subseteq \mathcal{L}$ and $\tilde{G}\mathcal{L}^\perp \subseteq \mathcal{L}^\perp$, then $\boldsymbol{x}$ is in optimal position if and only if (4.2) is valid and the configuration*

$$\boldsymbol{x}^o = \{x_1 - \mathrm{pr}_{\mathcal{L}} x_1, \ldots, x_n - \mathrm{pr}_{\mathcal{L}} x_n\}$$

is in optimal position with respect to $\tilde{G}$.

Proof. Necessity. If $\boldsymbol{x}$ is in optimal position, then $\boldsymbol{x}^o$ is in optimal position with respect to $\tilde{G}$, since the group of transformations

$$x \mapsto \tilde{g}x - \tilde{g}\mathrm{pr}_{\mathcal{L}} x\,, \quad \tilde{g} \in \tilde{G}$$

is a subgroup of G. Furthermore, the group of translations $x \mapsto x + l$, $l \in \mathcal{L}$ is a subgroup of G, whence (4.2) is valid.
Sufficiency. Since $x_i^o = x_i - \mathrm{pr}_{\mathcal{L}} x_i \in \mathcal{L}^\perp$, the conditions of Theorem 4 yield,

$$\begin{aligned}\frac{1}{n^2}\mathcal{I}(\boldsymbol{x}) &= \frac{1}{n^2}\sum_{1\le i<j\le n} \|\tilde{g}_i x_i^o - \tilde{g}_j x_j^o\|^2 \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left\| \tilde{g}_i \mathrm{pr}_{\mathcal{L}} x_i + l_i - \frac{1}{n}\sum_{j=1}^{n}(\tilde{g}_j \mathrm{pr}_{\mathcal{L}} x_j + l_j)\right\|^2 .\end{aligned}$$

If $\boldsymbol{x}^o$ is in optimal position with respect to $\tilde{G}$ and $\mathrm{pr}_{\mathcal{L}} x_i = b$, $1 \le i \le n$, then the inertia of $\tilde{\boldsymbol{g}}\boldsymbol{x}^o = \{\tilde{g}_1 x_1^o, \ldots, \tilde{g}_n x_n^o\}$ (the first summand) is minimal for $\tilde{g}_i = e$, $1 \le i \le n$.
The second summand is minimal if and only if

$$l_i = b - \tilde{g}_i \mathrm{pr}_{\mathcal{L}} x_i\,, \quad i = 1, \ldots, n$$

for any $b \in \mathcal{L}$, whence $l_i = 0$. Therefore, $\boldsymbol{x}$ is in optimal position. □

Theorem 5. *If the group G admits the decomposition (4.3), conditions of Theorem 4 are valid, and $\tilde{G}$ is a compact topological group, then, for each $\boldsymbol{x}$, there exist elements $g_1, \ldots, g_n \in G$ such that the configuration $\boldsymbol{g}\boldsymbol{x}$ is in optimal position.*

Proof. Theorem 4 gives optimal translations of the elements of $\boldsymbol{x}$. Therefore, consider the "centred" configuration $\boldsymbol{x}^o$ defined in that theorem. Then notice that the infimum

$$\inf_{\tilde{g}_1,\ldots,\tilde{g}_n \in \tilde{G}} \mathcal{I}(\tilde{\boldsymbol{g}}\boldsymbol{x}^o)$$

is taken over a compact set and the function $\mathcal{I}(\tilde{\boldsymbol{g}}\boldsymbol{x}^o)$ is continuous with respect to $\tilde{\boldsymbol{g}} = \{\tilde{g}_1, \ldots, \tilde{g}_n\} \subset \tilde{G}$. □

Let us now consider four particular applications of our theory.

First, Gower's and Ziezold's theory of mean landmark configurations ([11], [24], [26]) has to be mentioned. It was the starting point for our general theory. There ρ_G is a special case of the procrustes metric (without scale transformations). Gower and Ziezold studied k-tuples in the complex plane and showed that a configuration of such k-tuples is in optimal position when all centres of gravity coincide. Thus it suffices to determine the optimum rotations. Goodall [11] adapted this method for shape analysis of landmark configurations.

Our second example is the case of *convex* compact sets, which are described by support functions. Each n-tuple $K_1, \ldots, K_n$ of convex compact sets corresponds uniquely to the configuration $\boldsymbol{x} = \{s_{K_1}, \ldots, s_{K_n}\}$ in the family of support functions on the unit sphere $\mathbb{S}^{d-1}$.

The group of proper motions of sets corresponds to the group G acting on $L^2(\mathbb{S}^{d-1})$ as follows

$$g s_K(t) = \tilde{g} s_K(t) + \langle l, t \rangle \,, \quad t \in \mathbb{S}^{d-1}, \ l \in \mathbb{R}^d \,, \tag{4.4}$$

where $\tilde{g} s_K(t)$ is the support function of the set $\tilde{g}K$ obtained as a rotation of K. Thus, the group G admits the decomposition (4.3) with the space $\mathcal{L}$ of linear functions $\langle t, u \rangle$, $t \in \mathbb{S}^{d-1}$, for all $u \in \mathbb{R}^d$.

For the following the Steiner point (see [18, p. 203]) is important. It is defined by

$$s(K) = \frac{1}{b_d} \int\limits_{\mathbb{S}^{d-1}} t s_K(t) \mathrm{d}t$$

for a convex set K. Note that $s(K) \in K$ and $s(\{u\}) = u$.

Theorem 6. *The configuration $\{s_{K_1}, \ldots, s_{K_n}\}$ of support functions considered as elements of $L^2(\mathbb{S}^{d-1})$ is in optimal position with respect to the group G acting as in (4.4) if and only if the configuration $\{s_{K_1^o}, \ldots, s_{K_n^o}\}$ is in optimal position with respect to the group $\tilde{G}$, where $K_i^o = K_i - s(K_i)$, $i = 1, \ldots, n$.*

Proof. The support functions of all convex sets form a convex subset of $L^2(\mathbb{S}^{d-1})$. Note that the Theorems 2 and 4 are valid also if H is a convex subset of a Hilbert space. To apply Theorem 4, let us find the projection of $s_K(\cdot)$ onto the space $\mathcal{L}$ defined above. Since this projection belongs to $\mathcal{L}$, it is of the form $\langle t, u \rangle$ for some $u \in \mathbb{R}^d$.

Furthermore,

$$\int\limits_{\mathbb{S}^{d-1}} \langle t, u \rangle \langle t, v \rangle \mathrm{d}t = \int\limits_{\mathbb{S}^{d-1}} s_K(t) \langle t, v \rangle \mathrm{d}t \,,$$

for all $v \in \mathbb{R}^d$. The latter equation is equivalent to

$$\left\langle \int\limits_{\mathbb{S}^{d-1}} \langle t, u \rangle t \mathrm{d}t, v \right\rangle = \left\langle \int\limits_{\mathbb{S}^{d-1}} s_K(t) t \mathrm{d}t, v \right\rangle \,,$$

whence $u = s(\{u\}) = s(K)$. Now the statement follows directly from Theorem 4. □

Thus, in studies of optimal positions of convex sets, their Steiner points play the same role as the centres of gravity (centroids) for configurations of points.

The third example is the case of sets which are shrinkable with respect to a given reference point. (In many engineering applications the centre of gravity is chosen without long discussion as that reference point, and the particles have a shape such that they are shrinkable with respect to it.) In this case the optimum location is given a priori, and it is natural to assume that the sets are centred at the origin. Then it is only necessary to find the optimum rotations around o. Rotations clearly correspond to shifts of the radius-vector functions.

The fourth example is the case of general compact sets, which are described by their indicator functions. Here we can offer only a heuristic solution, which is given in section 5. A problem is here that there is not a unique characteristic point playing a role similar to that of the centre of gravity in Ziezold's case or of the Steiner point in the case of support functions. The following example shows that the centre of gravity is not this point.

Example. Let K_1 be the square with vertex points $P_1 = (2, 1)$, $P_2 = (2, -1)$, $P_3 = (0, 1)$ and $P_4 = (0, -1)$. Furthermore let K_2 be the triangle with vertices P_1, P_2 and the origin. We consider the corresponding indicator functions as elements of $L^2(\mathbb{R}^d)$. The squared norm of their difference is the Lebesgue measure of the symmetric difference between the corresponding sets. The minimum of the Lebesgue measure of the symmetric difference of compact sets congruent to K_1 and K_2 is obtained at the sets K_1 and K_2 with the vertices given above. But these two sets do not have the same centre of gravity. If K_2 is shifted in horizontal direction so that the centre of gravity of the shifted set coincides with that of K_1, then the Lebesgue measure of the symmetric difference of the sets increases. Note that the same value of the Lebesgue measure of the symmetric difference is also obtained for three other pairs of sets congruent to K_1 and K_2.

5. An Example

Figure 1 shows a sample $K_1, \ldots, K_{100}$ of planar (nearly polygonal) particles. These compact sets are profiles resulting from a planar section of a three-dimensional ceramic structure which can be considered a random packing of long and thin polyhedra, see [13]. In the original microscopic image these sets had random locations and orientations, which are without interest for the problem considered here. Therefore, they are considered as representatives of compact figures. Their forms are almost convex, so non-convex sets were replaced by their convex hulls.

We describe the profiles by their indicator functions and use the following heuristic algorithm, which is similar to the Gower algorithm, see [12] and [26].

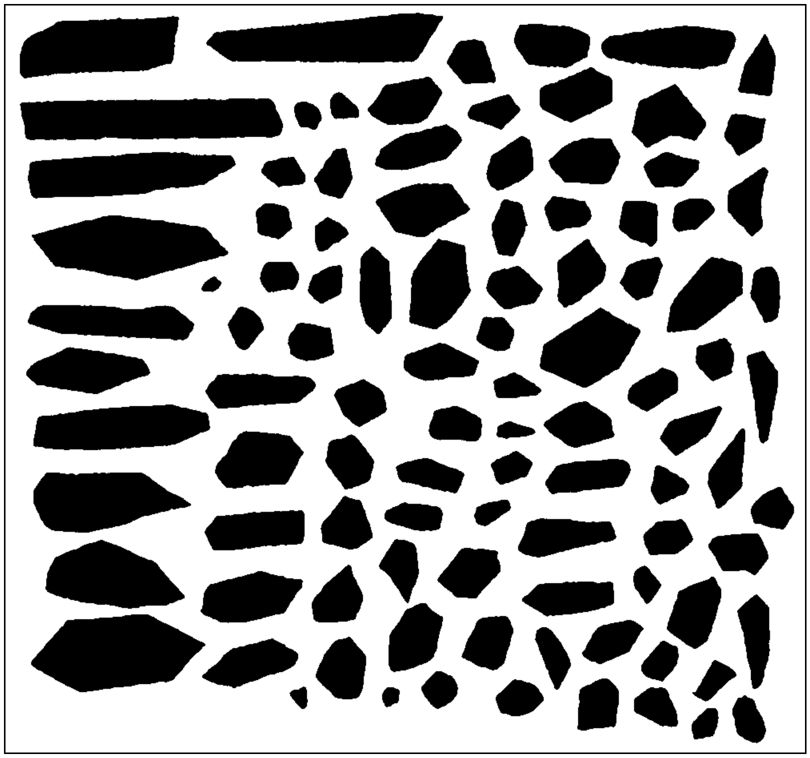


Figure 1: A sample of 100 compact sets.

Algorithm 3.

1. Determine the centroids (centres of gravity) c_i of the K_i and put $K_i^o = K_i - c_i$, $i = 1, \ldots, n$.

2. Choose as a starting function $f^{(0)}(\cdot) = 1_{K_1^o}(\cdot)$. Put $K_i^{(0)} = K_i^o$, $i = 1, \ldots, n$.

3. Rotate $K_i^{(j-1)}$ to obtain sets $K_i^{(j)}$ such that

$$\int_{\mathbb{R}^2} \left(1_{K_i^{(j)}}(x) - f^{(j-1)}(x)\right)^2 dx$$

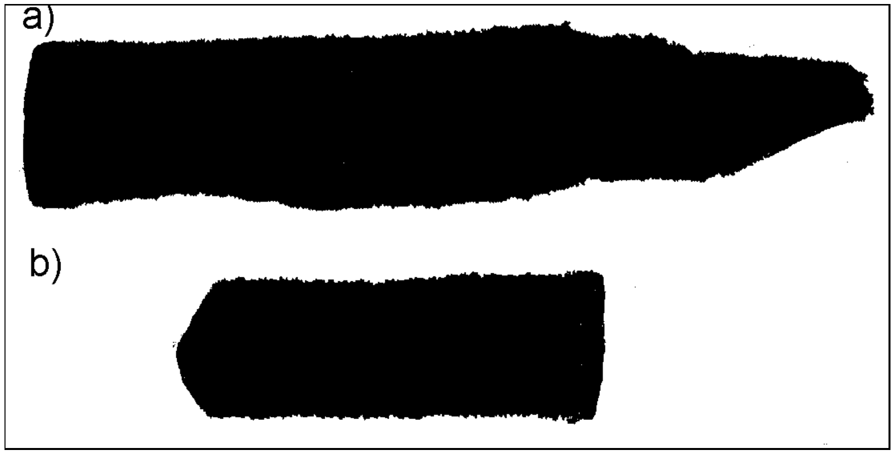
   becomes a minimum; $i = 1, 2, \ldots, n$.

4. Set $f^{(j)}(\cdot) = \frac{1}{n} \sum_{i=1}^{n} 1_{K_i^{(j)}}(\cdot)$.

   Stop the calculations, if $f^{(j)}$ is "close enough" to $f^{(j-1)}$ in the $L^2(\mathbb{R}^2)$ norm. Then write $\delta_i K_i^o = K_i^{(j)}$. Otherwise repeat step 3.

The resulting sets $K_i^{(j)}$ are obtained by rotations from the sets K_i^o, $i = 1, 2, \ldots, n$. Therefore we write $K_i^{(j)} = \delta_i K_i^o$.

Figure 2: Two starting figures for which Algorithm 3 was applied.



The final function $f^{(j)}$ in the algorithm is taken as an approximation of the relative mean of the family of indicator functions. For confirming the result, the algorithm can be started again with a different starting function $f^{(0)}$.

The procedure was used for two different starting functions, namely the indicator functions corresponding to the two figures shown in Figure 2.

The resulting means of indicator functions of rotated samples are shown in Figure 3 as level sets. Since the difference between these two functions is small, one may conclude that we are close to the theoretical result. The level sets of the coverage function clearly reflect the sizes of the profiles, even the size variability. Furthermore, various shape properties are visible: few of the profiles are very elongated, the smaller profiles tend to be more circular, and the majority of them are moderately elongated figures. These figures show important aspects of the mean behaviour of the profiles. Figure 3 also shows the corresponding Vorob'ev mean. For this example the set-theoretic median practically coincides with the Vorob'ev mean. It has an approximate elliptical shape.

We also analysed the same data by the support function approach. We found that the Steiner points of the sets K_i are close to the centres of gravity. Therefore we used the sets $\delta_i K_i^o$ as above i. e. we did not re-determine the optimal rotations corresponding to the support function approach. An approximation of the mean is then the Minkowski average

$$\bar{K} = n^{-1} \left(\delta_1 K_1^o \oplus \cdots \oplus \delta_n K_n^o \right) . \tag{5.1}$$

Figure 4 shows the Minkowski average of rotated sets $\delta_i K_i^o$ corresponding to Figure 3a. Note that the average on Figure 4 does not have the elliptical shape inherent to the contour lines of the mean indicator function. In contrast, it looks like the parallel set of a rectangle (rectangle $\oplus$ disk).

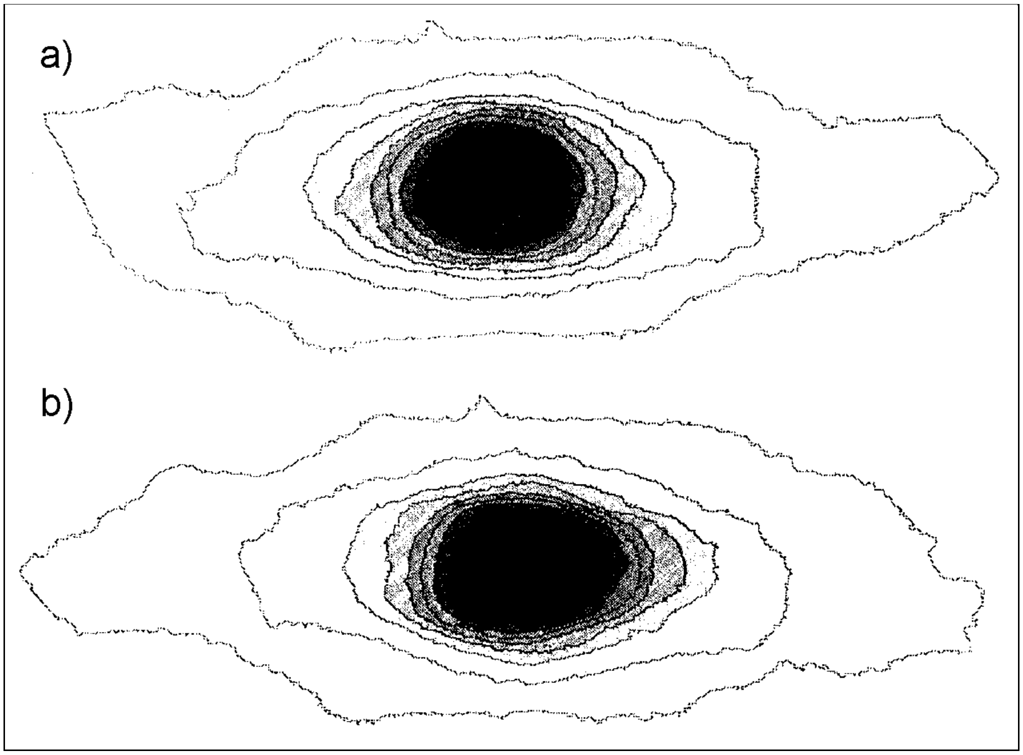


Figure 3: The resulting means of indicator functions transformed to optimal positions. The Vorob'ev mean is shown in black.

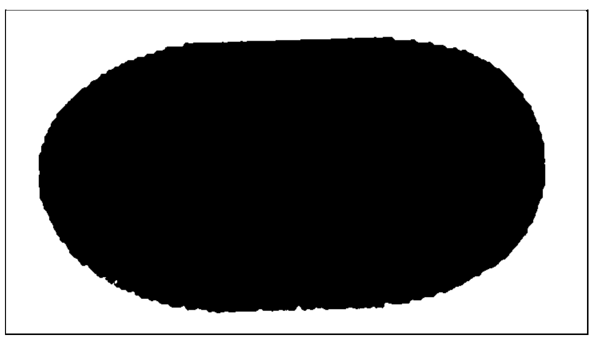


Figure 4: The Aumann expectation computed for the transformed sample.

6. Discussion

This paper presents first steps towards methods of exploratory data analysis for samples of particles or other irregular geometrical objects. It is shown how means, medians and variances can be determined. The indicator function mean produced by algorithm 3 may serve as a kind of empirical distribution function, since it shows important aspects of particle shape and size variability.

There are close relations and similarities to the problem of determining means of curves or functions [16] and [20]. However, instead of transformations which correspond to Euclidean motions there operations are used which synchronize the curves, for example, by dynamic time warping; see also [3].

An alternative approach to means of particles is using pseudo-landmarks and procrustes methods. A fixed large number of landmarks can be placed around each particle's outline. Then generalized procrustes analysis is carried out but with an additional step: permutation of the marks to best match each configuration to the current mean. Of course, since the theoretical approaches differ, also different results must be expected for the landmark and figure approach.

The problem of determining means suffers from possible non-uniqueness. The example in section 4 shows that for two particles several optimal positions are possible. This may lead to different means depending on the order in which the particles are included in the calculations. To avoid this danger the authors suggest to start the iteration algorithms with different starting figures. They hope that for real samples of particles, which are not mathematically symmetric figures, the non-uniqueness problem will be not so important. By the way, it cannot be avoided by using the landmark approach.

In this paper both size and shape are considered, size is not outperformed by scaling. Similarly as in the landmark case scaling is possible also for particles, but this seems not to make sense for the type of statistical analysis of particles the authors have in mind. However, scaling may lead to more robust results.

Appendix: Proof of Theorem 2

For the proof of Theorem 2 we use the following

Lemma. It is

$$\mathcal{I}(\boldsymbol{x}) = n(\mathcal{I}(\boldsymbol{x}, \overline{x}) - \mathcal{I}(\boldsymbol{x}))\,.$$

Proof. The Steiner theorem for moments of inertia implies that, for each i,

$$\sum_{j=1}^{n} \|x_i - x_j\|^2 = \sum_{j=1}^{n} \|x_j - \overline{x}\|^2 + n\|x_i - \overline{x}\|^2\,.$$

Summation for $1 \leq i \leq n$ yields

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \|x_i - x_j\|^2 = n\sum_{j=1}^{n} \|x_j - \overline{x}\|^2 + n\sum_{i=1}^{n} \|x_i - \overline{x}\|^2\,.$$

Thus, $2\mathcal{I}(\boldsymbol{x}) = 2n(\mathcal{I}(\boldsymbol{x}, \overline{x}) - \mathcal{I}(\boldsymbol{x}))$, and the Lemma is proved. □

Proof of Theorem 2. The proof of Theorem 2 is inspired by [24] and [26] result for point configurations in the complex finite-dimensional space.
Necessity. Let $\boldsymbol{x}$ be in optimal position. Then by the Lemma

$$n(\mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x})) = \mathcal{I}(\boldsymbol{x}) = \inf_{g_1,\dots,g_n \in G} \mathcal{I}(\boldsymbol{gx}) .$$

By the Lemma, the last term is equal to

$$\begin{aligned} n &\inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},\overline{gx}) - \mathcal{I}(\boldsymbol{gx})) \\ &= n \inf_{y\in H} \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},y) - \mathcal{I}(\boldsymbol{gx})) . \end{aligned}$$

Thus,

$$\begin{aligned} \mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x}) &= \inf_{y\in H} \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},y) - \mathcal{I}(\boldsymbol{gx})) \\ &\le \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},\overline{x}) - \mathcal{I}(\boldsymbol{gx})) . \end{aligned}$$

Furthermore,

$$\mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x}) \ \ge \ \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},\overline{x}) - \mathcal{I}(\boldsymbol{gx}))$$

since the left hand is obtained for unit group elements. Consequently,

$$\mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x}) = \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},\overline{x}) - \mathcal{I}(\boldsymbol{gx})) ,$$

and

$$\overline{x} \in \overline{x}_G .$$

Assume now that there exists an element $g \in G$ and an index i such that

$$\|gx_i - \overline{x}\| < \|x_i - \overline{x}\| .$$

Then, for $\boldsymbol{x}' = \{x_1, \dots, x_{i-1}, gx_i, x_{i+1}, \dots, x_n\}$ and $x'_i = gx_i$ we get

$$\begin{aligned} \mathcal{I}(\boldsymbol{x}',\overline{x'}) - \mathcal{I}(\boldsymbol{x}') &= \sum_{j=1}^{n} \|x'_j - \overline{x'}\|^2 \ \le \ \sum_{j=1}^{n} \|x'_j - \overline{x}\|^2 \\ &= \sum_{j=1}^{n} \|x_j - \overline{x}\|^2 + \|x'_i - \overline{x}\|^2 - \|x_i - \overline{x}\|^2 \\ &< \ \mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x}) . \end{aligned}$$

The Lemma yields $\mathcal{I}(\boldsymbol{x}') < \mathcal{I}(\boldsymbol{x})$, contrary to the assumption that $\boldsymbol{x}$ is in optimal position.

Sufficiency. Suppose that the conditions of Theorem 2 are valid. The Lemma and the condition $\overline{x} \in \overline{x}_G$ yield

$$\begin{aligned} \inf_{g_1,\dots,g_n\in G} \mathcal{I}(\boldsymbol{gx}) &= n \inf_{g_1,\dots,g_n\in G} \inf_{y\in H} (\mathcal{I}(\boldsymbol{gx},y) - \mathcal{I}(\boldsymbol{gx})) \\ &= n \inf_{g_1,\dots,g_n\in G} (\mathcal{I}(\boldsymbol{gx},\overline{x}) - \mathcal{I}(\boldsymbol{gx})) . \end{aligned}$$

Since each x_i and $\overline{x}$ are in optimal position, the last term is equal to $n(\mathcal{I}(\boldsymbol{x},\overline{x}) - \mathcal{I}(\boldsymbol{x}))$. Again by the Lemma, it is equal to $\mathcal{I}(\boldsymbol{x})$. □

Acknowledgements

The authors thank A. Seelig for constructive discussions, programming work and the calculations for Figures 3 and 4. The authors are grateful to W.F. Eddy, who kindly sent the thesis of L.A. Galway, and also to H. Ziezold, O.Yu. Vorob'ev and I. Dryden for their comments on earlier drafts of this paper.

References

1. R.J. Aumann, "Integrals of set-valued functions," *J. Math. Anal. Appl.* 12, pp. 1-12, 1965.
2. J.K. Beddow and T. Meloy, *Testing and Characterization of Powder and Fine Particles,* Heyden & Sons, London, 1980.
3. F.L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology,* Cambridge University Press, Cambridge, 1991.
4. G.E. Bredon, *Introduction to Compact Topological Groups,* Academic Press, London, 1972.
5. T.K. Carne, "The geometry of shape spaces," *Proc. Lond. Math. Soc.* 60, pp. 407–?, 1990.
6. I.L. Dryden and K.V. Mardia, "Theoretical and distributional aspects of shape analysis," In: *Probability Measures on Groups X* (ed. H. Heyer) Plenum Press, New York, pp. 95–116, 1991.
7. I.L. Dryden and K.V. Mardia, "Multivariate shape analysis," *Sankhya A* 55, pp. 460–480, 1993.
8. M. Fréchet, "Les elements aleatoires de nature quelconque dans un espace distancie," *Ann. Inst. H. Poincaré* 10, pp. 235–310, 1948.
9. L.A. Galway, *Statistical Analysis of Star-Shaped Sets,* Ph. D. Thesis, Carnegie-Mellon University, 1987.
10. C.R. Goodall, "Procrustes methods in the statistical analysis of shape," *J. Roy. Statist. Soc. B* 53, pp. 285–339, 1991.
11. J.C. Gower, "Generalized procrustes analysis," *Psychometrika* 40, pp. 33–51, 1975.
12. S. Hartmann, F. Mücklich, J. Ohser, W. Dreßler and G. Petzow, "Quantitative Charakterisierung von Si_3N_4-Gefügen durch räumliche Parameter." *Pract. Metallography, Special Issue* 24, 263–274, 1993.

13. A.E. Hawkins, *The Shape of Powder-Particle Outlines,* Research Studies Press, Taunton, and J. Wiley & Sons, New York, 1993.

14. H. Karcher, "Riemannian center of mass and mollifier smoothing," *Comm. Pure Appl. Math.* 30, pp. 509–541, 1977.

15. A. Kneip and T. Gasser, "Statistical tools to analyze data representing a sample of curves," *Ann. Statist.* 20, pp. 1266–1305, 1992.

16. H. Le and D.G. Kendall, "The Riemannian structure of euclidean shape space: a novel environment for statistics," *Ann. Statist.* 21, pp. 1225–1271, 1993.

17. K. Leichtweiß, *Konvexe Mengen,* VEB Deutscher Verlag der Wissenschaften, Berlin, 1980.

18. G. Matheron, *Random Sets and Integral Geometry,* Wiley, New York, 1975.

19. J.A. Rice and B. Silverman, "Estimating the mean and covariance structure non-parametrically when the data are curves," *J. Roy. Statist. Soc. B* 53, pp. 233–243, 1991.

20. D. Stoyan and H. Stoyan, *Fraktals, Random Shapes and Point Fields,* Wiley, Chichester, 1994.

21. R. Vitale, "An alternate formulation of mean value for random geometric figures," *J. Microscopy* 151, pp. 197–204, 1988.

22. O.Yu. Vorob'ev, *Srednemernoje Modelirovanie,* (*Mean-Measure Modelling*) Nauka, Moscow (in Russian), 1984.

23. H. Ziezold, "On expected figures and a strong law of large numbers for random elements in quasi-metric spaces," *Trans.* 7^{th} *Prague Conf. Inf. Th., Statist. Dec. Func., Random Processes (Prague, 1974).* Vol. A, Reidel, Dordrect, pp. 591–602, 1977.

24. H. Ziezold, "On expected figures in the plane, In: *Geobild'89, Math. Res. Series* 51, Akademie-Verlag, Berlin, pp. 105–110, 1989.

25. H. Ziezold, "Mean figures and mean shapes applied to biological figures and shape distributions in the plan, *Biom. J.* 36, pp. 491–510, 1994.