

Inner Products for Convex Bodies

David Bryant

*Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand
david.bryant@otago.ac.nz*

Petru Cioica-Licht

*Universität Duisburg-Essen, Fakultät für Mathematik, 45117 Essen, Germany
petru.cioica-licht@uni-due.de*

Lisa Orloff Clark

*School of Mathematics and Statistics, Victoria University, Wellington, New Zealand
lisa.clark@vuw.ac.nz*

Rachael Young

*Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand
rachael.gray.young@gmail.com*

Received: April 15, 2020

Accepted: December 22, 2020

We define a set inner product to be a function on pairs of convex bodies which is symmetric, Minkowski linear in each dimension, positive definite, and satisfies the natural analogue of the Cauchy-Schwartz inequality (which is not implied by the other conditions). We show that any set inner product can be embedded into an inner product space on the associated support functions, thereby extending fundamental results of Hörmander and Rådström. The set inner product provides a geometry on the space of convex bodies. We explore some of the properties of that geometry, and discuss an application of these ideas to the reconstruction of ancestral ecological niches in evolutionary biology.

Keywords: Inner product, convex body, Minkowski linear functionals, ecological niche.

2010 Mathematics Subject Classification: 52A20, 52A27, 05C05.

1. Introduction

A *convex body* is a closed, bounded, non-empty convex set. We use $\mathcal{K}$ or $\mathcal{K}_n$ to denote the set of convex bodies contained in $\mathbb{R}^n$. There is an extensive literature on convex bodies and their mathematics, with metric spaces and even Banach spaces defined on $\mathcal{K}$ (see [9] for a comprehensive review). Here we investigate an analog of a real *inner product space* defined on $\mathcal{K}$, motivated by applications in approximation and inference of convex bodies.

Our starting point is the definition of a *Minkowski linear functional* on $\mathcal{K}$: a function $f : \mathcal{K} \rightarrow \mathbb{R}$ satisfying $f(\alpha A + \beta B) = \alpha f(A) + \beta f(B)$ for all $A, B \in \mathcal{K}$ and $\alpha, \beta \geq 0$. A conventional real inner product is bilinear and positive definite. Replacing ‘bilinear’ with ‘Minkowski bilinear’ we obtain the following axioms for a set inner product $\langle \cdot, \cdot \rangle : \mathcal{K} \times \mathcal{K} \rightarrow \mathbb{R}$.

(A1) Symmetry: $\langle A, B \rangle = \langle B, A \rangle$ for all $A, B \in \mathcal{K}$.

(A2) Minkowski linearity with respect to the first variable: for all $\alpha, \beta \in \mathbb{R}_{\geq 0}$ and $A, B, C \in \mathcal{K}$,

$$\langle \alpha A + \beta B, C \rangle = \alpha \langle A, C \rangle + \beta \langle B, C \rangle. \quad (1)$$

(A3) Positive definite: $\langle A, A \rangle > 0$ when $A \neq \{0\}$.

The key difference between these axioms and those for a standard inner product is axiom (A2), as $\langle \alpha A, B \rangle = \alpha \langle A, B \rangle$ is only required for $\alpha \geq 0$. This small change makes a significant difference, including the fact that the Cauchy-Schwartz Inequality $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$ does not follow from (A1)–(A3), as we prove in the following section. We include the Cauchy-Schwartz inequality directly as a fourth axiom

(A4) For all $A, B \in \mathcal{K}$, $\langle A, B \rangle^2 \leq \langle A, A \rangle \langle B, B \rangle$ with equality if and only if $A = \lambda B$ for some $\lambda \in \mathbb{R}_{\geq 0}$.

Definition 1.1. A *set inner product* on $\mathcal{K}$ is a function $\langle \cdot, \cdot \rangle : \mathcal{K} \times \mathcal{K} \rightarrow \mathbb{R}$ satisfying (A1), (A2), (A3) and (A4).

Note that for any $a, b \in \mathbb{R}^n$ we have

$$\langle \{a\}, \{b\} \rangle + \langle -\{a\}, \{b\} \rangle = \langle \{a\} - \{a\}, \{b\} \rangle = \langle \{0\}, \{b\} \rangle = 0$$

so that $\langle \cdot, \cdot \rangle$ is bilinear when restricted to singletons. Thus the restriction of a set inner product to singleton sets is an inner product, called the *induced inner product*. We also note that for $A \in \mathcal{K}$, $-A := -1(A)$ is not the additive inverse of A and hence $\mathcal{K}$ is not a vector space. Note that replacing (A2) by the stronger version

(A2') For all $\alpha, \beta \in \mathbb{R}$ and $A, B, C \in \mathcal{K}$, $\langle \alpha A + \beta B, C \rangle = \alpha \langle A, C \rangle + \beta \langle B, C \rangle$

leads to a contradiction: since $A - A = -(A - A)$ and $\langle -B, B \rangle = -\langle B, B \rangle$ for all $A, B \in \mathcal{K}$ implies $\langle A - A, A - A \rangle = -\langle A - A, A - A \rangle = 0$ contradicting (A3).

Example 1.2. Let $\langle \cdot, \cdot \rangle : \mathcal{K}_1 \times \mathcal{K}_1 \rightarrow \mathbb{R}$ be the function on closed and bounded intervals of $\mathbb{R}$ given by

$$\langle [a, b], [c, d] \rangle = ac + bd.$$

Then $\langle \cdot, \cdot \rangle$ is a set inner product.

In this paper we investigate set inner products, with an eye to their potential application to the approximation and statistical inference of convex sets. Our main result is that set inner products are exactly inner products for the set of support functions (Theorem 2.1). This observation therefore extends the classical results of [4, 7] on embedding $\mathcal{K}$ in a Banach space. We conclude the second section by demonstrating that axiom (A4), the Cauchy-Schwartz inequality, is not implied by axioms (A1)–(A3).

On one hand Theorem 2.1 is a negative result: it shows that set inner products will not lead to fundamentally new mathematics which is not simply a consequence of standard functional analysis. On the other hand we obtain access to vast functional analysis toolkit to apply to the geometry and statistics of convex sets.

The geometry of $\mathcal{K}$ is the subject of the third section. We examine lines and subspaces in the set inner product space, and establish basic orthogonality results. We

make extensive use of Theorem 2.1 and classical results in convex analysis. Much is proved exclusively for an ℓ_2 set inner product and the extent to which these results can be generalised is not obvious.

In the final section, we show how the set inner product can be used to infer convex sets in comparative biology. Given an evolutionary tree and convex sets at the leaves (e.g. representing environmental niches) we use functional analysis techniques to infer niche geometries (convex sets) for ancestral species.

2. Characterization of set inner products

The *support function* of a convex body $A \in \mathcal{K}$ is defined

$$h_A : \mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \sup\{a \cdot x : a \in A\}.$$

where $\cdot$ is the usual dot product in $\mathbb{R}^n$. The properties of support functions are reviewed comprehensively in [9]. We let $\mathcal{H}$ denote the set of support functions for convex bodies in $\mathcal{K}$. Then $h \in \mathcal{H}$ if and only if it is *sublinear*: for all $x, y \in \mathbb{R}^n$ and $\alpha, \beta \geq 0$ we have

$$h(\alpha x + \beta y) \leq \alpha h(x) + \beta h(y). \quad (2)$$

Furthermore, for all $A, B \in \mathcal{K}$ and $\lambda \geq 0$ we have

$$h_{A+B} = h_A + h_B, \quad (3)$$

$$h_{\lambda A} = \lambda h_A. \quad (4)$$

and $A \subseteq B$ if and only if $h_A(x) \leq h_B(x)$ for all x . (5)

Let $\mathbb{S}^{n-1}$ denote the unit sphere in $\mathbb{R}^n$. Every function $h \in \mathcal{H}$ is determined by its restriction to $\mathbb{S}^{n-1}$; we denote this restriction by $\bar{h}$. Let $\mathcal{H}_S = \{\bar{h} : h \in \mathcal{H}\}$. The Hausdorff metric on $\mathcal{K}$

$$d_H(A, B) = \min\{r : A \subseteq B + r\mathbb{S}^{n-1} \text{ and } B \subseteq A + r\mathbb{S}^{n-1}\}$$

has an elegant re-expression in terms of support functions

$$d_H(A, B) = \|\bar{h}_A - \bar{h}_B\|_\infty = \sup\{|h_A(x) - h_B(x)| : x \in \mathbb{S}^{n-1}\},$$

see [4]. The set of convex bodies on $\mathbb{R}^n$ with the Hausdorff metric can be embedded naturally into a Banach space [4, 7]. See [8] for useful generalizations of these results. Schneider [9, pg. 45] observed that the span of $\mathcal{H}_S$ is dense in $\mathfrak{C}(\mathbb{S}^{n-1})$, the algebra of continuous functions on $\mathbb{S}^{n-1}$ (with respect to the L_∞ norm), and that every $f \in \mathfrak{C}(\mathbb{S}^{n-1})$ can be approximated arbitrarily closely by the difference $\bar{h}_A - \bar{h}_B$ of two support functions.

The Banach space embedding has proven a valuable tool for the study of random sets [5, 6] and also leads naturally to an L_p metric for $\mathcal{K}$:

$$d_p(A, B) = \|\bar{h}_A - \bar{h}_B\|_p \quad (6)$$

see [5, 13]. Arnold and Wellerding proposed a Sobolev distance on $\mathcal{K}$ [2]. Several authors have used the fact that this, like the L_2 distance given by (6), are inner product distances (e.g. [1, 10]). Here we attempt to put these results in a general setting.

Theorem 2.1. (Set inner product equivalence theorem) *Let $\langle \cdot, \cdot \rangle$ be a bivariate map on $\mathcal{K}$. Then $\langle \cdot, \cdot \rangle$ is a set inner product if and only if there is an inner product $\langle \cdot, \cdot \rangle_H$ on $\text{span}(\mathcal{H})$ such that*

$$\langle A, B \rangle = \langle h_A, h_B \rangle_H \quad \text{for all } A, B \in \mathcal{K}. \quad (7)$$

Proof. If there is such an inner product on $\text{span}(\mathcal{H})$ then the set inner product axioms for $\langle \cdot, \cdot \rangle$ follow directly from the definitions and properties of inner products and support functions.

For the converse, let $\langle \cdot, \cdot \rangle$ be a set inner product. Define $F : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ by $F(h_A, h_B) = \langle A, B \rangle$ for all A, B . We will use F to construct an inner product G on $\text{span}(\mathcal{H})$ which agrees with F on $\mathcal{H}$. As $\mathcal{H}$ is a convex cone, every element of $\text{span}(\mathcal{H})$ can be expressed as the difference of two elements in $\mathcal{H}$. Define a function G on $\text{span}(\mathcal{H}) \times \text{span}(\mathcal{H})$ by

$$G(f - g, h - k) = F(f, h) - F(f, k) - F(g, h) + F(g, k).$$

To see that G is well defined, suppose that $f_1, f_2, g_1, g_2, h_1, h_2, k_1, k_2 \in \mathcal{H}$. If $f_1 - g_1 = f_2 - g_2$ then

$$F(f_1, h_1) + F(g_2, h_1) = F(f_1 + g_2, h_1) = F(f_2 + g_1, h_1) = F(f_2, h_1) + F(g_1, h_1)$$

and

$$F(f_1, k_1) + F(g_2, k_1) = F(f_1 + g_2, k_1) = F(f_2 + g_1, k_1) = F(f_2, k_1) + F(g_1, k_1)$$

from which we obtain

$$\begin{aligned} F(f_1, h_1) - F(g_1, h_1) - F(f_1, k_1) + F(g_1, k_1) \\ = F(f_2, h_1) - F(g_2, h_1) - F(f_2, k_1) + F(g_2, k_1). \end{aligned}$$

Similarly, if $h_1 - k_1 = h_2 - k_2$ then

$$\begin{aligned} F(f_2, h_1) - F(g_2, h_1) - F(f_2, k_1) + F(g_2, k_1) \\ = F(f_2, h_2) - F(g_2, h_2) - F(f_2, k_2) + F(g_2, k_2). \end{aligned}$$

Hence if $f_1 - g_1 = f_2 - g_2$ and $h_1 - k_1 = h_2 - k_2$ then

$$\begin{aligned} F(f_1, h_1) - F(g_1, h_1) - F(f_1, k_1) + F(g_1, k_1) \\ = F(f_2, h_2) - F(g_2, h_2) - F(f_2, k_2) + F(g_2, k_2) \end{aligned}$$

We will show that G is an inner product on $\text{span}(\mathcal{H})$.

(i) As F is symmetric, so is G .

(ii) G is additive, since

$$\begin{aligned} G((f_1 - g_1) + (f_2 - g_2), h - k) \\ = F(f_1 + f_2, h) - F(f_1 + f_2, k) - F(g_1 + g_2, h) + F(g_1 + g_2, k) \\ = F(f_1, h) - F(f_1, k) - F(g_1, h) + F(g_1, k) + F(f_2, h) - F(f_2, k) - F(g_2, h) + F(g_2, k) \\ = G(f_1 - g_1, h - k) + G(f_2 - g_2, h - k). \end{aligned}$$

(iii) For all $\lambda \geq 0$

$$G(\lambda(f-g), h-k) = F(\lambda f, h) - F(\lambda f, k) - F(\lambda g, h) + F(\lambda g, k) = \lambda G(f-g, h-k)$$

and

$$\begin{aligned} G(-(f-g), h-k) &= G(g-f, h-k) \\ &= F(g, h) - F(g, k) - F(f, h) + F(f, k) = -G(f-g, h-k). \end{aligned}$$

Hence G is linear (and bilinear by symmetry).

(iv) For all $f, g \in \mathcal{H}$ there are A, B such that $f = h_A, g = h_B$. Then

$$\begin{aligned} G(f-g, f-g) &= F(f, f) - 2F(f, g) + F(g, g) \\ &= \langle A, A \rangle - 2\langle A, B \rangle + \langle B, B \rangle \\ &\geq \langle A, A \rangle - 2\sqrt{\langle A, A \rangle \langle B, B \rangle} + \langle B, B \rangle \quad \text{by (A4)} \\ &= (\sqrt{\langle A, A \rangle} - \sqrt{\langle B, B \rangle})^2 \geq 0. \end{aligned}$$

To see that G is positive definite notice

$$0 = G(f-g, f-g) \implies \langle A, A \rangle = \langle B, B \rangle = \langle A, B \rangle.$$

Axiom (A4) implies $A = \lambda B$ for some non-negative scalar λ and (A2) implies $\lambda = \lambda^2 = 1$ and hence $\lambda = 1$. Thus $f-g=0$. We have then that the map G is an inner product on $\text{span}(\mathcal{H})$. Furthermore, for any $A, B \in \mathcal{K}$ we have $\langle A, B \rangle = F(h_A, h_B) = G(h_A, h_B)$, giving (7). $\square$

Corollary 2.2. *Let $\langle \cdot, \cdot \rangle$ be a bivariate map on $\mathcal{K}$. Then $\langle \cdot, \cdot \rangle$ is a set inner product if and only if there is an inner product $\langle \cdot, \cdot \rangle_S$ on $\text{span}(\mathcal{H}_S)$ such that*

$$\langle A, B \rangle = \langle \bar{h}_A, \bar{h}_B \rangle_S \quad \text{for all } A, B \in \mathcal{K}.$$

Proof. If $\langle A, B \rangle = \langle \bar{h}_A, \bar{h}_B \rangle_S$ then axioms (A1)–(A4) follow directly from the definitions and properties of inner products and support functions. Conversely, if $\langle \cdot, \cdot \rangle$ is a set inner product then by the theorem there is an inner product $\langle \cdot, \cdot \rangle_H$ on $\text{span}(\mathcal{H})$ such that $\langle A, B \rangle = \langle h_A, h_B \rangle_H$ for all A, B . For each $\bar{f} \in \text{span}(\mathcal{H}_S)$ define $f \in \text{span}(\mathcal{H})$ by $f(x) = \|x\| \bar{f}\left(\frac{x}{\|x\|}\right)$ for all $x \neq 0$. Defining

$$\langle \bar{f}, \bar{g} \rangle_S = \langle f, g \rangle_H$$

gives the required inner product on $\text{span}(\mathcal{H}_S)$. $\square$

Example 2.3. Let $\phi(x)$ denote the standard Gaussian density on $\mathbb{R}^n$ and define

$$\langle\!\langle A, B \rangle\!\rangle = \int h_A(x) h_B(x) d\phi(x).$$

Then by Theorem 2.1 $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is a set inner product. Note that because $h_A(\lambda x) = \lambda h_A(x)$ for all $\lambda \geq 0$ we have that

$$\int h_A(x) h_B(x) d\phi(x) = \int_{\mathbb{S}^{n-1}} \bar{h}_A(x) \bar{h}_B(x) d\mu(x)$$

for the Haar measure μ on $\mathbb{S}^{n-1}$, the unit sphere in $\mathbb{R}^n$.

Suppose that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{R}^n$ and that we define the support functions with respect to this inner product in place of the dot product. If we pick an arbitrary inner product $\langle \cdot, \cdot \rangle_H$ on $\text{span}(\mathcal{H})$ then there is no guarantee that the resulting *set* inner product

$$\langle A, B \rangle_s = \langle h_A, h_B \rangle_H$$

will agree with the original inner product on singletons. That is, we could well have

$$\langle \{a\}, \{b\} \rangle_s = \langle h_{\{a\}}, h_{\{b\}} \rangle_H \neq \langle a, b \rangle$$

for some $a, b \in \mathbb{R}^n$. However for each such inner product $\langle \cdot, \cdot \rangle_H$ we can nevertheless find an invertible linear map $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that

$$\langle h_{\{a\}} \circ f, h_{\{b\}} \circ f \rangle_H = \langle a, b \rangle.$$

Example 2.4. The support functions for closed intervals on the real line are determined by two values, $h_{[a,b]}(1) = b$ and $h_{[a,b]}(-1) = -a$. Hence a generic set inner product on the real line can be written

$$\langle [a, b], [c, d] \rangle = \begin{bmatrix} b & -a \end{bmatrix} M \begin{bmatrix} d \\ -c \end{bmatrix}$$

where M is positive definite. The constraint that $\langle [a, a], [c, c] \rangle = ac$ is obtained by scaling M so that $\begin{bmatrix} 1 & -1 \end{bmatrix} M \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1$. $\square$

We made explicit use of the Cauchy-Schwartz inequality (A4) to prove that every set inner product corresponds to an inner product on the support functions. It is natural to ask whether this was strictly necessary, given that the inequality is implied by the axioms for standard inner products. The next construction gives a function which satisfies all of the set inner product axioms except the Cauchy-Schwartz inequality.

Proposition 2.5. *There is a bilinear function on $\mathcal{K}$ which satisfies (A1), (A2) and (A3) but not the Cauchy-Schwartz inequality (A4).*

Proof. Let $\{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$. Define $F : \mathcal{K} \times \mathcal{K} \rightarrow \mathbb{R}$ where

$$\begin{aligned} F(A, B) = & \frac{1}{8} (h_A(e_1)h_B(e_1) + h_A(e_2)h_B(e_2) + h_A(-e_1)h_B(-e_1) + h_A(-e_2)h_B(-e_2)) \\ & + [(h_A(e_1) + h_A(e_2) - h_A(e_1 + e_2)) (h_B(-e_1) + h_B(-e_2) - h_B(-e_1 - e_2))] \\ & + [(h_B(e_1) + h_B(e_2) - h_B(e_1 + e_2)) (h_A(-e_1) + h_A(-e_2) - h_A(-e_1 - e_2))] . \end{aligned}$$

Then F is symmetric and bilinear. For any $A \in \mathcal{K}$ we have

$$\begin{aligned} F(A, A) = & \frac{1}{8} (h_A(e_1)^2 + h_A(e_2)^2 + h_A(-e_1)^2 + h_A(-e_2)^2) + \\ & + 2 [(h_A(e_1) + h_A(e_2) - h_A(e_1 + e_2)) (h_A(-e_1) + h_A(-e_2) - h_A(-e_1 - e_2))] \geq 0. \end{aligned}$$

Now suppose $F(A, A) = 0$. Then since $h_A(e_1) + h_A(e_2) - h_A(e_1 + e_2) \geq 0$ and

$$h_A(-e_1) + h_A(-e_2) - h_A(-e_1 - e_2) \geq 0$$

we have $h_A(e_1) = h_A(e_2) = h_A(-e_1) = h_A(-e_2) = 0$.

Hence $h_A(\lambda e_i) = h_A(-\lambda e_i) = 0$ for $i = 1, 2$ and all $\lambda \geq 0$ and so for all $x \in \mathbb{R}^2$,

$$h_A(x) = h_A(x_1 e_1 + x_2 e_2) \leq h_A(x_1 e_1) + h_A(x_2 e_2) \leq 0.$$

By (5), $A \subseteq \{0\}$, giving equality as A is non-empty.

We have shown that F satisfies the first four axioms of a set inner product. We now show that it fails the last axiom. Let A be the unit circle in $\mathbb{R}^2$ and B be the triangle through the points $(0, 0)$, $(1, 0)$, $(0, 1)$. Then

$$\begin{aligned} h_A(e_1) &= h_A(e_2) = h_A(-e_1) = h_A(-e_2) = 1; \\ h_A(e_1 + e_2) &= h_A(-e_1 - e_2) = \sqrt{2}; \\ h_B(e_1) &= h_B(e_2) = h_B(e_1 + e_2) = 1; \\ h_B(-e_1) &= h_B(-e_2) = h_B(-e_1 - e_2) = 0. \end{aligned}$$

We then have

$$\begin{aligned} 2F(A, B) - 2\sqrt{F(A, A)F(B, B)} &= (\sqrt{F(A, A)} - \sqrt{F(B, B)})^2 - F(A, A) + \\ &+ 2F(A, B) - F(B, B) \geq -F(A, A) + 2F(A, B) - F(B, B) = 6\sqrt{2} - \frac{33}{4} \approx 0.235 \end{aligned}$$

so that the Cauchy-Schwarz inequality fails. $\square$

3. Lines and subspaces in $\mathcal{K}$

Having established that set inner products are essentially standard inner products (for support functions) we explore implications for a geometry of the space of convex bodies.

Definition 3.1. A subset $\mathcal{A} \subseteq \mathcal{K}$ is an *affine subspace* if $\{h_A : A \in \mathcal{A}\}$ equals the intersection of $\mathcal{H}$ and an affine subspace of $\text{span}(\mathcal{H})$. A *subspace* of $\mathcal{K}$ is an affine subspace which includes the zero-set $\{0\}$.



Figure 3.1: The line segment in $\mathcal{K}_2$ connecting a circle to a square.

Example 3.2. Figure 3.1 depicts elements on the line segment

$$\{A \in \mathcal{K}_2 : h_A = \alpha h_S + (1 - \alpha)h_C \text{ for some } \alpha \in [0, 1]\}$$

connecting a square S and a circle C in $\mathcal{K}_2$. Figure 3.2 depicts the two dimensional subspace of $\mathcal{K}_2$ spanned by a circle C and a triangle T ,

$$\{A \in \mathcal{K}_2 : h_A = \alpha h_C + \beta h_T \text{ for some } \alpha, \beta \geq 0\}.$$

Note that the first example is bounded and the second example is unbounded. Any subspace which contains a non-zero convex body A also contains λA for all $\lambda \geq 0$, so that the subspace is unbounded. The same need not hold for affine subspaces.

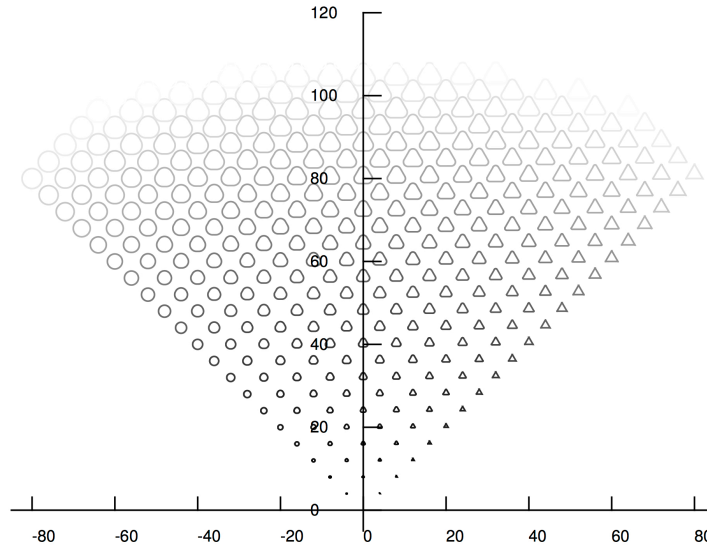


Figure 3.2: The two dimensional subspace of $\mathcal{K}_2$ spanned by a circle at $(-40, 40)$ and an equilateral triangle at $(40, 40)$.

We say that a subset ℓ of $\mathcal{K}$ is a *line* if the corresponding set of support functions $\{h_A : A \in \ell\}$ equals the intersection of a line with $\mathcal{H}$. Because of the intersection with $\mathcal{H}$, lines can have endpoints and subspaces can be bounded subsets.

Proposition 3.3. *Every line $\ell \subseteq \mathcal{K}$ is one of three types (where $A, B \in \mathcal{K}$, $x \in \mathbb{R}^n$) :*

- (1) Translations: $\{A + tx : t \in \mathbb{R}\}$
- (2) Rays: $\{tA + B : t \geq 0\}$
- (3) Segments: $\{tA + (1 - t)B : 0 \leq t \leq 1\}$.

Proof. Suppose that $A, B \in \ell$, so that $X \in \ell$ if and only if

$$h_X \in \{h_B + t(h_A - h_B) : t \in \mathbb{R}\}.$$

We consider three cases.

Case 1: $h_A - h_B = h_{\{x\}}$ for some $x \in \mathbb{R}^n$. Then ℓ is a translation.

Case 2: $h_A - h_B = h_C$ for some non-singleton $C \in \mathcal{K}$ so $\ell = \{h_B + th_C : t \in \mathbb{R}\}$. As C is not a singleton, h_C is not linear and there are x, y such that $h_C(x + y) < h_C(x) + h_C(y)$. We then have that if t is sufficiently negative,

$$h_B(x + y) + th_C(x + y) > h_B(x) + th_C(x) + h_B(y) + th_C(y)$$

so that the set $\{t : h_B + th_C \in \mathcal{H}\}$ is bounded below.

Let $\lambda = \min\{t : h_B + th_C \in \mathcal{H}\}$. Then $X \in \ell$ if and only if

$$h_X \in \{(h_B + \lambda h_C) + th_C : t \geq 0\}$$

so that ℓ is a ray.

Case 3: $h_A - h_B \notin \mathcal{H}$. Let $f = h_A - h_B$. As $f \notin \mathcal{H}$ there is $x, y \in \mathbb{R}^n$ such that $f(x) + f(y) - f(x+y) < 0$. Hence if

$$t > \frac{h_B(x) + h_B(y) - h_B(x+y)}{f(x+y) - f(x) - f(y)}$$

then $h_B + t(h_A - h_B) \notin \mathcal{H}$. Thus the line has an endpoint at A or at some point on the other side of A from B . The same holds in the opposite direction. $\square$

The characterisation of endpoints was considered (in three dimensions) by [12] and is closely related to Minkowski summands and decompositions. For $K, L \in \mathcal{K}$ we say that L is a *summand* of K if there is $M \in \mathcal{K}$ such that $L + M = K$. Schneider [9, pg. 157] shows that this is equivalent to being able to cover K with translates of L contained in K .

Proposition 3.4. *Suppose $A, B \in \mathcal{K}$. Then A is an endpoint on the line through A and B if and only if for all $\epsilon \in (0, 1)$, the convex body ϵB is not a summand of A .*

Proof. Suppose there is C such that $\epsilon B + C = A$. Let $t = \frac{-\epsilon}{1-\epsilon}$. Then $t < 0$ and

$$(1-t)h_A + th_B = \frac{1}{1-\epsilon}h_C$$

so $\frac{1}{1-\epsilon}C$ is on the line through A and B on the opposite side of A from B .

Conversely, if there is $h_C \in \mathcal{H}$ and $t < 0$ such that $h_C = (1-t)h_A + th_B$ then

$$\frac{-t}{1-t}h_B + \frac{1}{1-t}h_C = \frac{-t}{1-t}h_B + \frac{1-t}{1-t}h_A + \frac{t}{1-t}h_B = h_A,$$

so $(\frac{-t}{1-t})B + (\frac{1}{1-t})C = A$, where $(\frac{-t}{1-t})B$ is a summand of C . $\square$

Which convex bodies can be endpoints? Intuitively, these are convex bodies on the boundary of $\mathcal{H}$.

Proposition 3.5. *Every convex body is the endpoint of some line.*

Proof. Suppose that $A \in \mathcal{K}$ and let x be an exposed point of A (that is, there is a supporting hyperplane H of A such that $\{x\} = H \cap A$). Let B be any line segment containing x and contained in H . Then any translation of a scaled version of B which contains x will also contain other points in H which are therefore not in A . It follows that ϵB does not slide freely around A for any $\epsilon \in (0, 1)$, and A is an endpoint of the line through A and B . $\square$

A convex body which has no summands (other than scaled and translated versions of itself) is said to be *indecomposable*. These convex bodies are exactly those which are endpoints of *every* line they are contained in. In $\mathbb{R}^2$ the indecomposable bodies are the line segments and triangles [11], while for higher dimensions, the indecomposable bodies form a dense subset of $\mathcal{K}_n$ with respect to the Hausdorff metric (see [9, pg.165-6] for details).

4. Geometry of the set inner product

Let $\langle \cdot, \cdot \rangle$ be a set inner product on $\mathcal{K}_n$, where we write

$$\langle a, b \rangle = \langle \{a\}, \{b\} \rangle$$

for the induced inner product.

For any convex body $A \in \mathcal{K}$, the map $x \mapsto \langle A, \{x\} \rangle$ is a continuous linear functional. By the Riesz-Fréchet theorem there is a unique element $k_A \in \mathbb{R}^n$ such that

$$\langle A, \{x\} \rangle = \langle k_A, x \rangle$$

for all $x \in \mathbb{R}^n$. We say that k_A is the *center* of A .

Example 4.1. Recall the L_2 set inner product

$$\langle\langle A, B \rangle\rangle = \int h_A(x) h_B(x) d\phi(x)$$

defined in Example 2.3. The center of a set for this set inner product is given by the *Steiner point*

$$s_A = \int z h_A(z) d\phi(z)$$

since for all $x \in \mathbb{R}^n$ we have $h_{\{x\}}(z) = \langle x, z \rangle$ and

$$\langle\langle A, \{x\} \rangle\rangle = \int \langle z, x \rangle h_A(z) d\phi(z) = \left\langle \int z h_A(z) d\phi(z), x \right\rangle = \langle s_A, x \rangle.$$

The Steiner point for a convex body always lies in the relative interior of the set [9, pg. 43] so for this set inner product, $k_A \in A$. Furthermore, for any set inner product we have for all $x \in \mathbb{R}^n$ that

$$\langle k_{\{a\}}, x \rangle = \langle \{a\}, \{x\} \rangle = \langle a, x \rangle$$

so $k_{\{a\}} = a$. However in general, we do not have that $k_A \in A$.

Example 4.2. Let $\langle \cdot, \cdot \rangle$ be a set inner product for closed intervals on the real line, where

$$\langle [a, b], [c, d] \rangle = \begin{bmatrix} b & -a \end{bmatrix} M \begin{bmatrix} d \\ -c \end{bmatrix} \quad \text{and} \quad \langle [a, a], [c, c] \rangle = ac.$$

Then for $A = [a, b]$ we have $k_A = \begin{bmatrix} b & -a \end{bmatrix} M \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so that for all $x \in \mathbb{R}$,

$$k_A x = \begin{bmatrix} b & -a \end{bmatrix} M \begin{bmatrix} x \\ -x \end{bmatrix} = \langle [a, b], [x, x] \rangle.$$

As an example, let $M = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$, so M is positive definite and $\begin{bmatrix} 1 & -1 \end{bmatrix} M \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1$.

However $k_A = [b \ -a] \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2a - b$. Hence if $A = [0, 1]$ then $k_A \notin A$.

We say that a convex body $A \in \mathcal{K}$ is *centered* if $k_A = 0$, which holds if and only if $\langle A, \{x\} \rangle = 0$ for all x . We write $A_0 = A - \{k_A\} := A + \{-k_A\}$, noting that for all x ,

$$\langle A_0, \{x\} \rangle = \langle A, \{x\} \rangle + \langle -k_A, \{x\} \rangle = 0$$

so that A_0 is centered.

Proposition 4.3. *Suppose that $A_0 = A - k_A$ and $B_0 = B - k_B$. Then*

$$\langle A, B \rangle = \langle A_0, B_0 \rangle + \langle k_A, k_B \rangle.$$

Proof. We have

$$\begin{aligned} \langle A, B \rangle &= \langle A_0 + k_A, B_0 + k_B \rangle \\ &= \langle A_0, B_0 \rangle + \langle A_0, k_B \rangle + \langle k_A, B_0 \rangle + \langle k_A, k_B \rangle \\ &= \langle A_0, B_0 \rangle + \langle A, k_B \rangle - \langle k_A, k_B \rangle + \langle k_A, B \rangle - \langle k_A, k_B \rangle + \langle k_A, k_B \rangle \\ &= \langle A_0, B_0 \rangle + \langle k_A, k_B \rangle. \end{aligned}$$

For all $x \in \mathbb{R}^n$ we have $\langle a, x \rangle = \langle \{a\}, \{x\} \rangle = \langle k_a, x \rangle$. Thus $k_a - a = 0$. $\square$

For the remainder of this section we explore properties of the L_2 set inner product $\langle\langle \cdot, \cdot \rangle\rangle$ introduced in Example 2.3. It is not clear which of these results can be generalised.

Proposition 4.4. *Suppose that $\langle \cdot, \cdot \rangle$ equals the L_2 inner product*

$$\langle\langle A, B \rangle\rangle = \int h_A(x)h_B(x)d\phi(x).$$

- (1) For all $A \in \mathcal{K}$, $k_A \in A$.
- (2) If A and B are centered then $\langle A, B \rangle = 0$ if and only if $A = \{0\}$ or $B = \{0\}$.
- (3) If $0 \in A \subseteq B$ then $\langle A, A \rangle \leq \langle B, B \rangle$.

Proof. (1) See [9, pg. 43].

(2) Suppose by way of contradiction that neither A nor B equal $\{0\}$. As $0 \in A \cap B$ we have that h_A and h_B are both non-negative. We will show that there is x such that $h_A(x)h_B(x) > 0$, implying that

$$\int h_A(x)h_B(x)d\phi(x) > 0.$$

If there is x such that the strict half-space $\{y : \langle x, y \rangle > 0\}$ intersects both A and B then

$$h_A(x)h_B(x) \geq \langle a, x \rangle \langle b, x \rangle > 0.$$

Otherwise, A and B are both contained in a line $\{tv : v \in \mathbb{R}\}$ and they only intersect at 0.

Without loss of generality, $A \subset \{tv : t \geq 0\}$. However then $h_A(x)h_{\{v\}}(x) \geq 0$ for all x and $h_A(v)h_{\{v\}}(v) > 0$, implying

$$\langle A, \{v\} \rangle = \int h_A(x)h_{\{v\}}(x)d\phi(x) > 0 \quad \text{and } k_A \neq 0.$$

(3) As $0 \in A$ and $A \subseteq B$ we have $0 \leq h_A(x) \leq h_B(x)$ for all x . $\square$

We write $A - A = A + (-1)A = \{a_1 - a_2 : a_1, a_2 \in A\}$.

Lemma 4.5. *If $A \cap B \neq \emptyset$ then $(A \cup B) - (A \cup B) \subseteq (A - A) + (B - B)$.*

Proof. Since $0 \in A - A$ and $0 \in B - B$ we have $(A - A) \cup (B - B) \subseteq (A - A) + (B - B)$. Thus if $x, y \in A$ or $x, y \in B$ then $x - y \in (A - A) + (B - B)$.

Suppose $z \in A \cap B$. If $x \in A$ and $y \in B$ then $x - z \in A - A$ and $z - y \in B - B$ so $x - y = (x - z) + (z - y) \in (A - A) + (B - B)$. $\square$

A *diversity* (X, δ) is an extension of the concept of a metric space to include comparisons of more than two points at a time [3]. Here X is a set and δ is a function on finite subsets of X satisfying

(D1) $\delta(A) \geq 0$ and $\delta(A) = 0 \Leftrightarrow |A| \leq 1$

(D2) For all A, B, C with $B \neq \emptyset$, $\delta(A \cup B) + \delta(B \cup C) \geq \delta(A \cup C)$.

We note that (D1) and (D2) imply that δ is monotonic with respect to inclusion: if $A \subseteq B$ then $\delta(A) \leq \delta(B)$.

Proposition 4.6. *For any finite $A \subseteq X$ define*

$$\delta(A) = \sqrt{\langle \text{conv}(A - A), \text{conv}(A - A) \rangle}.$$

Then δ is a diversity.

Proof. We check the axioms. If $A = \{x\}$ then $A - A = \{0\}$ and $\delta(A) = 0$.

If $A \subseteq B$ then $A - A \subseteq B - B$ and $\text{conv}(A - A) \subseteq \text{conv}(B - B)$. As $0 \in A - A$ we have $\delta(A) \leq \delta(B)$ by Proposition 4.4 part 3. That is δ is monotonic.

If $a, b \in A$ and $a \neq b$ then $\text{conv}(A - A) \neq \{0\}$ and $\delta(A) \geq \delta(\{a, b\}) > 0$. Suppose that $z \in A \cap B$. By the lemma we have

$$(A \cup B) - (A \cup B) \subseteq (A - A) + (B - B)$$

which implies

$$\text{conv}((A \cup B) - (A \cup B)) \subseteq \text{conv}(A - A) + \text{conv}(B - B).$$

Hence

$$\begin{aligned} \delta(A \cup B) &= \langle \text{conv}(A \cup B - A \cup B), \text{conv}(A \cup B - A \cup B) \rangle^{1/2} \\ &\leq \langle \text{conv}(A - A) + \text{conv}(B - B), \text{conv}(A - A) + \text{conv}(B - B) \rangle^{1/2} \\ &\leq \langle \text{conv}(A - A), \text{conv}(A - A) \rangle^{1/2} + \langle \text{conv}(B - B), \text{conv}(B - B) \rangle^{1/2} \\ &= \delta(A) + \delta(B). \end{aligned}$$

This, and the fact that δ is monotonic, implies (D2). $\square$

5. Projecting convex bodies onto phylogenies

In this final section we present an application of set inner products to computational biology. In ecology, a *niche* can be defined as the set of environments in which an organism or species can survive/prosper. One can model the niche of a homogeneous population as a convex subset within some abstract space of environmental parameters. We are interested in studying how niches change over time.

A *phylogeny* is a tree T with leaves (degree one nodes) labelled bijectively by species and internal nodes corresponding to ancestral species. We assume that the tree is *binary*: every internal node has degree three. The phylogeny is model of the history of speciations or splits given rise to the species in the sample.

The general problem we consider is the following: suppose we have inferred niche information for each of the observed (leaf) species; can we infer how those niche differences evolved over time, and can we infer anything about the niches occupied by ancestral species?

A comprehensive answer to these questions will require models of random and directed change based on stochastic processes together with assessment of model and statistical uncertainty. Much of the original motivation for this paper was setting up the machinery for such models. However we will see that set inner products already provide a direct estimate based on a Steiner-tree problem for ancestral niches.

We first consider a special case. Consider N species all children of a single ancestral node. This is often called a *star tree* (Fig. 5.1). Let $\langle \cdot, \cdot \rangle$ be an inner product. We measure the distance between convex sets using the associated norm:

$$d(A, B) = (\langle h_A - h_B, h_A - h_B \rangle)^{1/2}.$$

Each niche is modelled by a convex body. The aim is to determine a convex body X for the ancestral node which minimizes the total *squared parsimony* length of the tree:

$$\sum_{i=1}^N d(A_i, X)^2.$$

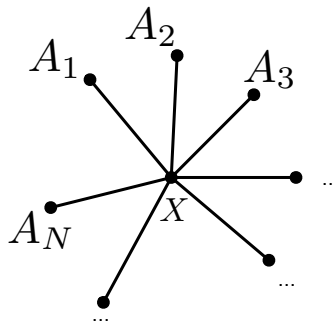


Figure 5.1: A star tree

This problem can be solved for a general inner product space.

Proposition 5.1. *The convex body X which minimizes $\sum_{i=1}^N d(A_i, X)^2$ is given by*

$$X = \frac{1}{N} \sum_{i=1}^N A_i.$$

Proof. This is a direct consequence of two observations. First, that $g = \frac{1}{N} \sum_{i=1}^N h_{A_i}$ is the unique minimizer of

$$\min_f \sum_{i=1}^N \|g - h_{A_i}\|^2 = \min_f \sum_{i=1}^N \langle g - h_{A_i}, g - h_{A_i} \rangle$$

on the completion of $(\text{span}(\mathcal{H}), \langle \cdot, \cdot \rangle)$. Second, that $g = \frac{1}{N} \sum_{i=1}^N h_{A_i}$ is in the convex hull of a set of support functions, so is itself a support function for a convex body X . $\square$

Let T be a phylogeny with N leaves and let $V(T)$, $L(T)$, and $E(T)$ denote the sets of nodes, leaves and edges of T . Let $\chi : L(T) \rightarrow \mathcal{K}_n$ be a map from the leaves of T to the set of convex bodies. An *extension* of χ is a function $\hat{\chi} : V(T) \rightarrow \mathcal{K}_n$ such that $\hat{\chi}(x) = \chi(x)$ for leaves $x \in L(T)$. The (squared parsimony) length of $\hat{\chi}$ is defined

$$\ell(\hat{\chi}) = \sum_{\{x,y\} \in E(T)} d(\hat{\chi}(x), \hat{\chi}(y))^2. \quad (8)$$

The aim is to find an extension $\hat{\chi}$ of minimum length.

Theorem 5.2. *There exists a unique extension $\hat{\chi}$ of minimum length. For each $v \in V(T)$ the convex body $\hat{\chi}(v)$ can be written as a non-negative linear combination*

$$\sum_{x \in L(T)} \lambda_{vx} \chi(x)$$

of convex bodies at the leaves (and with $\lambda_{xx} = 1$ for all $x \in L(T)$). Furthermore, the coefficients λ_{vx} do not depend on the underlying set inner product, nor on the map χ .

Proof. Suppose that $\hat{\chi}$ is an extension with minimum length. Let v be any node in $V(T) \setminus L(T)$ and let a, b, c be its neighbours. We have from Proposition 5.1 that

$$\hat{\chi}(v) = \frac{1}{3}(\hat{\chi}(a) + \hat{\chi}(b) + \hat{\chi}(c)). \quad (9)$$

That is $\hat{\chi}$ satisfies a type of Laplacian system. We will show by construction that these equations can be simultaneously satisfied for all internal nodes.

Fix a leaf $x \in L(T)$ and direct all edges in T away from x . For any node $v \neq x$ we let $p_x(v)$ denote the neighbor of v on the path from v to x . For each node v define $\alpha_v = 0$ if v is a leaf and otherwise let $\alpha_v = (3 - \alpha_a - \alpha_b)^{-1}$, where a and b are the two nodes with $p_x(a) = p_x(b) = v$. For each internal node v let λ_{vx} be the product of α_u over all internal nodes u on the path from v to x . We repeat this process for all x .

First note that if v is any internal node, a, b, c are its neighbors and $c = p_x(v)$ then

$$\begin{aligned} \frac{1}{3}(\lambda_{ax} + \lambda_{bx} + \lambda_{cx}) &= \frac{1}{3}(\alpha_a \alpha_v + \alpha_b \alpha_v + 1) \lambda_{cx} \\ &= \frac{1}{3} \left(\frac{\alpha_a}{3 - \alpha_a - \alpha_b} + \frac{\alpha_b}{3 - \alpha_a - \alpha_b} + 1 \right) \lambda_{cx} = \alpha_v \lambda_{cx} = \lambda_{vx}. \end{aligned}$$

As this holds for all x we have

$$\hat{\chi}(v) = \frac{1}{3}(\hat{\chi}(a) + \hat{\chi}(b) + \hat{\chi}(c)).$$

To prove uniqueness, suppose that $\chi(y) = \{0\}$ for all y . Fix x and direct the edges in the tree as above. We claim that if $\hat{\chi}$ satisfies (9) then $\hat{\chi}(v) = \{0\}$ for all v . To this end, we show first that $\hat{\chi}(v) = \alpha_v \hat{\chi}(p(v))$ for all nodes $v \neq x$. The result holds trivially for all leaves x . Suppose that v has neighbors a, b, c and that $p(v) = c$ and that the $\hat{\chi}(a) = \alpha_a \hat{\chi}(v)$ and $\hat{\chi}(b) = \alpha_b \hat{\chi}(v)$. We have

$$\hat{\chi}(v) = \frac{1}{3}(\hat{\chi}(a) + \hat{\chi}(b) + \hat{\chi}(c)) = \frac{1}{3}(\alpha_a \hat{\chi}(v) + \alpha_b \hat{\chi}(v) + \hat{\chi}(c))$$

so

$$\hat{\chi}(v) = \frac{1}{3 - \alpha_a - \alpha_b} \hat{\chi}(c) = \alpha_v \hat{\chi}(c).$$

If v is adjacent to x then $\hat{\chi}(v) = \alpha_v 0$, from which we have $\hat{\chi}(v) = 0$ for all v . $\square$

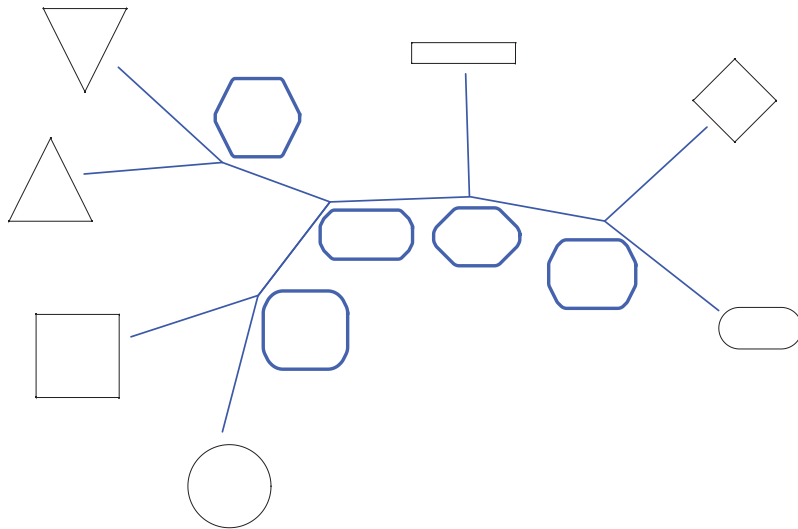


Figure 5.2: A tree with seven leaves, each labelled by a convex shape (in black). The shapes minimizing the total length of the tree (Eq. 8) are drawn in bold next to the corresponding internal nodes.

The proof of Theorem 5.2 leads directly to a time optimal $O(n^2)$ time algorithm for computing the coefficients λ_{vx} . Figure 5.2 shows the result of applying the algorithm to a set of simple 2D convex bodies on a tree with seven leaves. The original shapes label the leaves of the tree while the shapes giving a minimal length label the nodes in the interior of the tree. Source code for determining the coefficients λ_{vx} is available from the corresponding author.

References

- [1] R. Arnold: *Zur L^2 -Bestapproximation eines konvexen Körpers durch einen bewegten konvexen Körper*, Monatshefte f. Mathematik 108/4 (1989) 277–293.
- [2] R. Arnold, A. Wellerding: *On the Sobolev distance of convex bodies*, Aequationes Mathematicae 44 (1992) 72–83.
- [3] D. Bryant, P. F. Tupper: *Hyperconvexity and tight-span theory for diversities*, Advances in Mathematics 231/6 (2012) 3172–3198.
- [4] L. Hörmander: *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe*, Arkiv f. Matematik 3/2 (1955) 181–186.
- [5] D. E. McClure, R. A. Vitale: *Polygonal approximation of plane convex bodies*, J. Math. Analysis Appl. 51/2 (1975) 326–358.
- [6] I. S. Molchanov: *Theory of Random Sets*, Springer, Berlin (2005).
- [7] H. Rådström: *An embedding theorem for spaces of convex sets*, Proc. Amer. Math. Soc. 3/1 (1952) 165–169.
- [8] K. D. Schmidt: *Embedding theorems for classes of convex sets*, Acta Appl. Mathematica 5/3 (1986) 209–237.
- [9] R. Schneider: *Convex Bodies: the Brunn-Minkowski Theory*, Cambridge University Press, Cambridge (2014).
- [10] G. C. Shephard: *A pre-Hilbert space consisting of classes of convex sets*, Israel J. Math. 4/1 (1966) 1–10.
- [11] R. Silverman: *Decomposition of plane convex sets II*, Pacific J. Math. 47/2 (1973) 521–530.
- [12] P. Vincensini: *Sur le prolongement des series lineaires de corps convexes. Applications*, Rend. Circ. Mat. Palermo 60/1 (1936) 361–372.
- [13] R. A. Vitale: *L_p metrics for compact, convex sets*, J. Approximation Theory 45/3 (1985) 280–287.