A Course
on
Convex Geometry

Daniel Hug, Wolfgang Weil
University of Karlsruhe
revised version 2009/2010

February 5, 2010

Preface

The following notes were written before and during the course on *Convex Geometry* which was held at the University of Karlsruhe in the winter term 2002/2003. Although this was the first course on this topic which was given in English, the material presented was based on previous courses in German which have been given several times, mostly in summer terms. In comparison with these previous courses, the standard program was complemented by sections on surface area measures and projection functions as well as by a short chapter on integral geometric formulas. The idea here was to lay the basis for later courses on *Stochastic Geometry*, *Integral Geometry* etc., which usually follow in a subsequent term.

The exercises at the end of each section contain all the weekly problems which were handed out during the course and discussed in the weakly exercise session. Moreover, I have included a few additional exercises (some of which are more difficult) and even some hard or even unsolved problems. The list of exercises and problems is far from being complete, in fact the number decreases in the later sections due to the lack of time while preparing these notes.

I thank Matthias Heveling and Markus Kiderlen for reading the manuscript and giving hints for corrections and improvements.

Karlsruhe, February 2003 Wolfgang Weil

During repetitions of the course in 2003/2004 and 2005/2006 a number of misprints and small errors have been detected. They are corrected in the current version. Also, additional material and further exercises have been added.

Karlsruhe, October 2007 Wolfgang Weil

During the courses in 2008/2009 (by D. Hug) and 2009/2010 (by W. Weil) these lecture notes have been revised and extended again. Also, some pictures have been included.

Karlsruhe, October 2009 Daniel Hug and Wolfgang Weil

Contents

Bibliography

[Al] A.D. Aleksandrov, *Konvexe Polyeder.* Akademie-Verlag, Berlin 1958.

[Ba] I.J. Bakelman, *Convex Analysis and Nonlinear Geometric Elliptic Equations.* Springer Berlin et al. 1994.

[Bar] A. Barvinok, *A Course in Convexity.* AMS, Providence, RI 2002.

[Be] R.V. Benson, *Euclidean Geometry and Convexity.* McGraw-Hill, New York 1966.

[BKOS] M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf, *Computational Geometry – Algorithms and Applications.* Springer, Berlin, Second Revised Edition, 2000.

[Bl] W. Blaschke, *Kreis und Kugel.* 2. Aufl., Walter der Gruyter, Berlin 1956.

[Boe] K. Böröczky Jr., *Finite Packing and Covering.* Cambridge University Press, Cambridge 2004.

[BY] J-D. Boissonnat, M. Yvinec, *Algorithmic Geometry.* Cambridge University Press 1998 (English edition).

[BMS] V. Boltyanski, H. Martini, P.S. Soltan, *Excursions into Combinatorial Geometry.* Springer, Berlin et al. 1997.

[BF] T. Bonnesen, W. Fenchel, *Theorie der konvexen Körper.* Springer, Berlin 1934.

[BMP] P. Brass, W. Moser, J. Pach, *Research Problems in Discrete Geometry.* Springer, New York 2005.

[Bro] A. Brønsted, *An Introduction to Convex Polytopes.* Springer, Berlin et al. 1983.

[BZ] Y.D. Burago, V.A. Zalgaller, *Geometric Inequalities.* Springer, Berlin et al. 1988.

[Bu] H. Busemann, *Convex Surfaces.* Interscience Publ., New York 1958.

[Ed] H. Edelsbrunner, *Algorithms in Combinatorial Geometry.* Springer, Berlin 1987.

[Eg] H.G. Eggleston, *Convexity.* Cambridge Univ. Press, London et al. 1958.

[Ew] G. Ewald, *Combinatorial Convexity and Algebraic Geometry.* Springer, New York et al. 1996.

[FT] L. Fejes Tóth, *Lagerungen in der Ebene, auf der Kugel und im Raum.* 2. verb. u. erw. Aufl. Springer, Berlin 1972.

[Ga] R.J. Gardner, *Geometric Tomography.* Cambridge Univ. Press, Cambridge 1995. Revised 2nd edition 2006.

[Go] J.E. Goodman, J. O'Rourke, *Handbook of Discrete and Computational Geometry.* CRC Press, Boca Raton, 1997.

[Gro] H. Groemer, *Geometric Applications of Fourier Series and Spherical Harmonics.* Cambridge Univ. Press 1996

[Gr] P.M. Gruber, *Convex and Discrete Geometry.* Grundlehren der mathematischen Wissenschaften Bd. 336, Springer, Berlin 2007.

[Gru] B. Grünbaum, *Convex Polytopes.* Interscience Publ., London et al. 1967. 2. ed. (prepared by Volker Kaibel). Springer, New York, 2003.

[Gru2] B. Grnbaum, G.C. Shepard, *Tilings and Patterns: an Introduction.* Freeman, New York, 1989.

[Ha1] H. Hadwiger, *Altes und Neues über konvexe Körper.* Birkhäuser, Basel et al. 1955.

[Ha2] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie.* Springer, Berlin et al. 1957.

[HaDe] H. Hadwiger, H. Debrunner, V. Klee, *Combinatorial Geometry in the Plane.* Holt, Rinehart and Winston, New York, 1964.

[Ho] L. Hörmander, *Notions of Convexity.* Birkhäuser, Basel et al. 1994.

[JoTh] M. Joswig, Th. Theobald, *Algorithmische Geometrie: polyedrische und algebraische Methoden.* Vieweg, Wiesbaden, 2008.

[KW] L. Kelly, M.L. Weiss, *Geometry and Convexity.* Wiley/Interscience Publ., New York et al. 1979.

[Kl] R. Klein, *Algorithmische Geometrie*. Addison-Wesley-Longman, Bonn 1997.

[Ko1] A. Koldobsky, *Fourier Analysis in Convex Geometry*. Mathematical Surveys and Monographs, American Mathematical Society, Providence RI 2005.

[KY] A. Koldobsky, V. Yaskin, *The Interface between Convex Geometry and Harmonic Analysis.* CBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence RI 2008.

[Le1] K. Leichtweiß, *Konvexe Mengen.* Springer, Berlin et al. 1980.

[Le2] K. Leichtweiß, *Affine Geometry of Convex Bodies.* J.A. Barth, Heidelberg et al. 1998.

[Le] M. Leppmeier, *Kugelpackungen von Kepler bis heute. Eine Einführung für Schüler, Studenten und Lehrer.* Vieweg, Braunschweig 1997.

[Ly] L.A. Lyusternik, *Convex Figures and Polyhedra.* Dover Publ., New York 1963.

[Ma] J.T. Marti, *Konvexe Analysis.* Birkhäuser, Basel et al. 1977.

[Mat] J. Matoušek, *Lectures on Discrete Geometry*. Graduate Texts in Mathematics, Vol. 212, Springer, New York, 2002.

[MS] P. McMullen, G.C. Shephard, *Convex Polytopes and the Upper Bound Conjecture.* Cambridge Univ. Press, Cambridge 1971.

[PA] J. Pach, P.K. Agarval, *Combinatorial Geometry.* Wiley-Interscience Series, Whiley, New York, 1995.

[Roc] R.T. Rockafellar, *Convex Analysis.* Princeton Univ. Press, Princeton 1970.

[Rog] C.A. Rogers, *Packing and Covering.* Cambridge University Press, Cambridge 1964.

[OR] J. O'Rourke, *Computational Geometry in C.* Cambridge University Press, Cambridge 1994.

[S] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory.* Cambridge Univ. Press, Cambridge 1993.

[SW] R. Schneider, W. Weil, *Integralgeometrie.* Teubner, Stuttgart 1992.

[SW2] R. Schneider, W. Weil, *Stochastische Geometrie.* Teubner, Stuttgart 2000.

[SW3] R. Schneider, W. Weil, *Stochastic and Integral Geometry.* Springer, Berlin 2008.

[StW] J. Stoer, Ch. Witzgall, *Convexity and Optimization in Finite Dimensions I.* Springer, Berlin et al. 1970.

[Ths] R. R. Thomas, *Lectures on Geometric Combinatorics.* Lecture Notes, University of Washington, Seattle 2004.

[Th] A.C. Thompson, *Minkowski Geometry.* Cambridge Univ. Press, Cambridge 1996.

[Va] F.A. Valentine, *Convex Sets.* McGraw-Hill, New York 1964. Deutsche Fassung: *Konvexe Mengen.* BI, Mannheim 1968.

[We] R. Webster, *Convexity.* Oxford Univ. Press, New York 1964.

[Zi] G.M. Ziegler, *Lectures on Polytopes.* Springer, Berlin et al. 1995. Revised 6th printing 2006.

[Zo1] C. Zong, *Strange Phenomena in Convex and Discrete Geometry.* Springer, New York 1996.

[Zo2] C. Zong, *Sphere Packings.* Springer, New York 1999.

[Zo3] C. Zong, *The Cube: a Window to Convex and Discrete Geometry.* Cambridge University Press, Cambridg, 2006.

Introduction

Convexity is an elementary property of a set in a real (or complex) vector space V. A set $A \subset V$ is convex if it contains all the segments joining any two points of A, i.e. if $x, y \in A$ and $\alpha \in [0, 1]$ implies that $\alpha x + (1 - \alpha)y \in A$. This simple algebraic property has surprisingly many and far-reaching consequences of geometric nature, but it also has topological consequences (if V carries a compatible topology) as well as analytical ones (if the notion of convexity is extended to real functions via their graphs). The interplay between convex sets and functions turns out to be particularly fruitful. Results on convex sets and functions play a central role in many mathematical fields, in particular in functional analysis, in optimization theory and in stochastic geometry.

During this course, we shall concentrate on convex sets in $\mathbb{R}^n$ as the prototype of a finite dimensional real vector space. In infinite dimensional spaces often other methods have to be used and different types of problems occur. Here, we concentrate on the classical part of convexity. Starting with convex sets and their basic properties (in Chapter 1), we briefly discuss convex functions (in Chapter 2), and then come (in Chapter 3) to the theory of convex bodies (compact convex sets). Our goal here is to present the essential parts of the Brunn-Minkowski theory (mixed volumes, quermassintegrals, Minkowski inequalities, in particular the isoperimetric inequality) as well as some more special topics (surface area measures, projection functions). In the last chapter, we will shortly discuss selected basic formulas from integral geometry. If time permits we will discuss symmetrization of convex sets and functions in an additional chapter.

The course starts rather elementary. Apart from a good knowledge of linear algebra (and, in Chapter 2, analysis) no deeper knowledge of other fields is required. Later we will occasionally use results from functional analysis, in some parts, we require some familiarity with topological notions and, more importantly, we use some concepts and results from measure theory.

Preliminaries and notations

Throughout the course we work in n-dimensional Euclidean space $\mathbb{R}^n$. Elements of $\mathbb{R}^n$ are denoted by lower case letters like $x, y, \ldots, a, b, \ldots$, scalars by greek letters $\alpha, \beta, \ldots$ and (real) functions by $f, g, \ldots$ We identify the vector space structure and the affine structure of $\mathbb{R}^n$, i.e. we do not distinguish between vectors and points. The coordinates of a point $x \in \mathbb{R}^n$ are used only occasionally, therefore we indicate them as $x = (x^{(1)}, \ldots, x^{(n)})$. We equip $\mathbb{R}^n$ with its usual topology generated by the standard scalar product

$$\langle x, y \rangle := x^{(1)}y^{(1)} + \cdots + x^{(n)}y^{(n)}, \quad x, y \in \mathbb{R}^n,$$

and the corresponding Euclidean norm

$$\|x\| := ((x^{(1)})^2 + \cdots + (x^{(n)})^2)^{1/2}, \quad x \in \mathbb{R}^n.$$

By B^n we denote the unit ball,

$$B^n := \{x \in \mathbb{R}^n : \|x\| \leq 1\},$$

and by

$$S^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$$

the unit sphere. Sometimes, we also make use of the Euclidean metric $d(x, y) := \|x - y\|$, $x, y \in \mathbb{R}^n$. Sometimes it is convenient to write $\frac{x}{\alpha}$ instead of $\frac{1}{\alpha}x$, for $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

Convex sets in $\mathbb{R}^1$ are not very exciting (they are open, closed or half-open, bounded or unbounded intervalls), usually results on convex sets are only interesting for $n \geq 2$. In some situations, results only make sense, if $n \geq 2$, although we shall not emphasize this in all cases. As a rule, $A, B, \ldots$ denote general (convex or nonconvex) sets, $K, L, \ldots$ will be used for compact convex sets (convex bodies) and $P, Q, \ldots$ for (convex) polytopes.

A number of notations will be used frequently, without further explanations:

$\operatorname{lin} A$	linear hull of A
$\operatorname{aff} A$	affine hull of A
$\dim A$	dimension of A (= dimension of $\operatorname{aff} A$)
$\operatorname{int} A$	interior of A
$\operatorname{rel\,int} A$	relative interior of A (interior w.r.t. $\operatorname{aff} A$)
$\operatorname{cl} A$	closure of A
$\operatorname{bd} A$	boundary of A
$\operatorname{rel\,bd} A$	relative boundary of A

If f is a function on $\mathbb{R}^n$ with values in $\mathbb{R}$ or in the extended real line $[-\infty, \infty]$ and if A is a subset of the latter, we frequently abbreviate the set $\{x \in \mathbb{R}^n : f(x) \in A\}$ by $\{f \in A\}$. Hyperplanes $E \subset \mathbb{R}^n$ are therefore shortly written as $E = \{f = \alpha\}$, where f is a linear form, $f \neq 0$, and $\alpha \in \mathbb{R}$ (note that this representation is not unique). The corresponding closed half-spaces generated by E are then $\{f \geq \alpha\}$ and $\{f \leq \alpha\}$, and the open half-spaces are $\{f > \alpha\}$ and $\{f < \alpha\}$.

The symbol $\subset$ always includes the case of equality. The abbreviation w.l.o.g. means 'without loss of generality' and is used sometimes to reduce the argument to a special case. The logical symbols $\forall$ (for all) and $\exists$ (exists) are occasionally used in formulas. $\square$ denotes the end of a proof. Finally, we write $|A|$ for the cardinality of a set A.

Each section is complemented by a number of exercises. Some are very easy, but most require a bit of work. Those which are more challenging than it appears from the first look are marked by $*$. Occasionally, problems have been included which are either very difficult to solve or even unsolved up to now. They are indicated by P.

Chapter 1

Convex sets

1.1 Algebraic properties

The definition of a convex set requires just the structure of $\mathbb{R}^n$ as a vector space. In particular, it should be compared with the notions of a linear and an affine subspace.

Definition. A set $A \subset \mathbb{R}^n$ is *convex*, if $\alpha x + (1-\alpha)y \in A$ for all $x, y \in A$ and $\alpha \in [0,1]$.

Examples. (1) The simplest convex sets (apart from the points) are the segments. We denote by

$$[x,y] := \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$$

the *closed segment* between x and y, $x, y \in \mathbb{R}^n$. Similarly,

$$(x,y) := \{\alpha x + (1-\alpha)y : \alpha \in (0,1)\}$$

is the *open segment* and we define half-open segments $(x,y]$ and $[x,y)$ in an analogous way.

(2) Other trivial examples are the affine flats in $\mathbb{R}^n$.

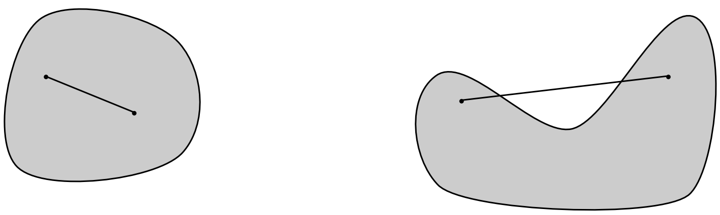
(3) If $\{f = \alpha\}$ ($f \neq 0$ a linear form, $\alpha \in \mathbb{R}$) is the representation of a hyperplane, the *open half-spaces* $\{f < \alpha\}, \{f > \alpha\}$ and the *closed half-spaces* $\{f \leq \alpha\}, \{f \geq \alpha\}$ are convex.

(4) Further convex sets are the *balls*

$$B(r) := \{x \in \mathbb{R}^n : \|x\| \leq r\}, \quad r \geq 0,$$

and their translates.

(5) Another convex set and a nonconvex set:



Let $k \in \mathbb{N}$, let $x_1, \ldots, x_k \in \mathbb{R}^n$, and let $\alpha_1, \ldots, \alpha_k \in [0,1]$ with $\alpha_1 + \ldots \alpha_k = 1$, then $\alpha_1 x_1 + \cdots + \alpha_k x_k$ is called a *convex combination* of the points $x_1, \ldots, x_k$.

Theorem 1.1.1. *A set $A \subset \mathbb{R}^n$ is convex, if and only if all convex combinations of points in A lie in A.*

Proof. Taking $k = 2$, we see that the condition on the convex combinations implies convexity.

For the other direction, assume A is convex and $k \in \mathbb{N}$. We use induction on k.

For $k = 1$, the assertion is trivially fulfilled.

For the step from $k-1$ to k, $k \geq 2$, assume $x_1, \ldots, x_k \in A$ and $\alpha_1, \ldots, \alpha_k \in [0,1]$ with $\alpha_1 + \ldots \alpha_k = 1$. We may assume $\alpha_i \neq 0$, $i = 1, \ldots, k$, and define

$$\beta_i := \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_{k-1}}, \quad i = 1, \ldots, k-1,$$

hence $\beta_i \in [0,1]$ and $\beta_1 + \ldots + \beta_{k-1} = 1$. By the induction hypothesis, $\beta_1 x_1 + \ldots + \beta_{k-1} x_{k-1} \in A$, and by the convexity

$$\sum_{i=1}^{k} \alpha_i x_i = \left(\sum_{i=1}^{k-1} \alpha_i \right) \left(\sum_{i=1}^{k-1} \beta_i x_i \right) + \left(1 - \sum_{i=1}^{k-1} \alpha_i \right) x_k \in A.$$

□

If $\{A_i : i \in I\}$ is an arbitrary family of convex sets (in $\mathbb{R}^n$), then the intersection $\bigcap_{i \in I} A_i$ is convex. In particular, for a given set $A \subset \mathbb{R}^n$, the intersection of all convex sets containing A is convex, it is called the *convex hull* $\operatorname{conv} A$ of A.

The following theorem shows that $\operatorname{conv} A$ is the set of all convex combinations of points in A.

Theorem 1.1.2. *For $A \subset \mathbb{R}^n$,*

$$\operatorname{conv} A = \left\{ \sum_{i=1}^{k} \alpha_i x_i : k \in \mathbb{N}, x_1, \ldots, x_k \in A, \alpha_1, \ldots, \alpha_k \in [0,1], \sum_{i=1}^{k} \alpha_i = 1 \right\}.$$

Proof. Let B denote the set on the right-hand side. If C is a convex set containing A, Theorem 1.1.1 implies $B \subset C$. Hence, we get $B \subset \operatorname{conv} A$.

On the other hand, the set B is convex, since

$$\begin{aligned} &\beta(\alpha_1 x_1 + \cdots + \alpha_k x_k) + (1-\beta)(\gamma_1 y_1 + \cdots + \gamma_m y_m) \\ &\qquad = \beta\alpha_1 x_1 + \cdots + \beta\alpha_k x_k + (1-\beta)\gamma_1 y_1 + \cdots + (1-\beta)\gamma_m y_m, \end{aligned}$$

for $x_i, y_j \in A$ and coefficients $\beta, \alpha_i, \gamma_j \in [0,1]$ with $\alpha_1 + \ldots + \alpha_k = 1$ and $\gamma_1 + \ldots + \gamma_m = 1$, and

$$\beta\alpha_1 + \cdots + \beta\alpha_k + (1-\beta)\gamma_1 + \cdots + (1-\beta)\gamma_m = \beta + (1-\beta) = 1.$$

Since B contains A, we get $\operatorname{conv} A \subset B$. □

Remarks. (1) Trivially, A is convex, if and only if $A = \operatorname{conv} A$.

(2) Later, in Section 1.2, we will give an improved version of Theorem 1.1.2 (CARATHEODORY's theorem), where the number k of points used in the representation of $\operatorname{conv} A$ is bounded by $n+1$.

Definition. For sets $A, B \subset \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$, we put

$$\alpha A + \beta B := \{\alpha x + \beta y : x \in A, y \in B\}.$$

The set $\alpha A + \beta B$ is called a *linear combination* of the sets A, B, the operation $+$ is called *vector addition*. Special cases get special names:

$A + B$	the *sum set*
$A + x$ (the case $B = \{x\}$)	a *translate* of A
αA	the *multiple* of A
$\alpha A + x$ (for $\alpha \geq 0$)	a *homothetic image* of A
$-A := (-1)A$	the *reflection* of A (in the origin)
$A - B := A + (-B)$	the *difference* of A and B

Remarks. (1) If A, B are convex and $\alpha, \beta \in \mathbb{R}$, then $\alpha A + \beta B$ is convex.

(2) In general, the relations $A + A = 2A$ and $A - A = \{0\}$ are **wrong**. For a convex set A and $\alpha, \beta \geq 0$, we have $\alpha A + \beta A = (\alpha + \beta)A$. The latter property characterizes convexity of a set A.

We next show that convexity is preserved by affine transformations.

Theorem 1.1.3. *Let $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ be convex and $f : \mathbb{R}^n \to \mathbb{R}^m$ affine. Then*

$$f(A) := \{f(x) : x \in A\}$$

and

$$f^{-1}(B) := \{x \in \mathbb{R}^n : f(x) \in B\}$$

are convex.

Proof. Both assertions follow from

$$\alpha f(x) + (1 - \alpha)f(y) = f(\alpha x + (1 - \alpha)y).$$

□

Corollary 1.1.4. *The projection of a convex set onto an affine subspace is convex.*

The converse is obviously false, a shell bounded by two concentric balls is not convex but has convex projections.

Definition. (a) The intersection of finitely many closed half-spaces is called a *polyhedral set*.

(b) The convex hull of finitely many points $x_1, \ldots, x_k \in \mathbb{R}^n$ is called a (convex) *polytope* P.

(c) The convex hull of affinely independent points is called a *simplex*, an r*-simplex* is the convex hull of $r+1$ affinely independent points.

Intuitively speaking, the vertices of a polytope P form a minimal set of points from P which generate the polytope. A precise definition is the following.

Definition. A point x of a polytope P is called a vertex of P, if $P \setminus \{x\}$ is convex. The set of all vertices of P is denoted by $\operatorname{vert} P$.

Theorem 1.1.5. *Let P be a polytope in $\mathbb{R}^n$, and let $x_1, \ldots, x_k \in \mathbb{R}^n$ be distinct points.*
(a) *If $P = \operatorname{conv}\{x_1, \ldots, x_k\}$, then x_1 is a vertex of P, if and only if $x_1 \notin \operatorname{conv}\{x_2, \ldots, x_k\}$.*
(b) *P is the convex hull of its vertices.*

Proof. **(a)** If x_1 is a vertex of P, then $x_1 \notin P \setminus \{x_1\}$. Since $P \setminus \{x_1\}$ is convex, we get $\operatorname{conv}\{x_2, \ldots, x_k\} \subset P \setminus \{x_1\}$, and hence $x_1 \notin \operatorname{conv}\{x_2, \ldots, x_k\}$.

Conversely, assume that $x_1 \notin \operatorname{conv}\{x_2, \ldots, x_k\}$. If x_1 is not a vertex of P, then there exist distinct points $a, b \in P \setminus \{x_1\}$ and $\lambda \in (0,1)$ such that $x_1 = (1-\lambda)a + \lambda b$. Hence there exist $k \in \mathbb{N}$, $\mu_1, \ldots, \mu_k \in [0,1]$ and $\tau_1, \ldots, \tau_k \in [0,1]$ with $\mu_1 + \ldots + \mu_k = 1$ and $\tau_1 + \ldots + \tau_k = 1$ such that $\mu_1, \tau_1 \neq 1$ and

$$a = \sum_{i=1}^{k} \mu_i x_i, \qquad b = \sum_{i=1}^{k} \tau_i x_i.$$

Thus we get

$$x_1 = \sum_{i=1}^{k} \left((1-\lambda)\mu_i + \lambda\tau_i\right) x_i,$$

from which it follows that

$$x_1 = \sum_{i=2}^{k} \frac{(1-\lambda)\mu_i + \lambda\tau_i}{1-(1-\lambda)\mu_1 - \lambda\tau_1} x_i, \tag{1.1}$$

where $(1-\lambda)\mu_1 + \lambda\tau_1 \neq 1$ and the right-hand side of (1.1) is a convex combination of $x_2, \ldots, x_k$, a contradiction.

(b) Using (a), we can successively remove points from $\{x_1, \ldots, x_k\}$ which are not vertices without changing the convex hull. Moreover, if $x \notin \{x_1, \ldots, x_k\}$ and x is a vertex of P, then $P = \operatorname{conv}\{x, x_1, \ldots, x_k\}$ implies that $x \notin \operatorname{conv}\{x_1, \ldots, x_k\} = P$, a contradiction. □

Remarks. (1) A polyhedral set is closed and convex. Polytopes, as convex hulls of finite sets, are closed and bounded, hence compact. We discuss these topological questions in more generality in Section 1.3.

(2) For a polytope P, Theorem 1.1.5 shows that $P = \operatorname{conv} \operatorname{vert} P$. This is a special case of MINKOWSKI's theorem, which is proved in Section 1.5.

(3) Polyhedral sets and polytopes are somehow dual notions. We shall see later in Section 1.4 that the set of polytopes coincides with the set of bounded polyhedral sets.

(4) The polytope property is preserved by the usual operations. In particular, if P, Q are polytopes, then the following sets are polytopes as well:

• $\mathrm{conv}\,(P \cup Q)$,
• $P \cap Q$,
• $\alpha P + \beta Q$, for $\alpha, \beta \in \mathbb{R}$,
• $f(P)$, for an affine map $f : \mathbb{R}^n \to \mathbb{R}^m$.

Here, only the second assertion is not straight-forward. The proof that $P \cap Q$ is a polytope will follow later for instance from the mentioned connection between polytopes and bounded polyhedral sets.

(5) If P is the convex hull of affinely independent points $x_0, \dots, x_r$, then each x_i is a vertex of P, i.e. P is an r-simplex. An r-simplex P has dimension $\dim P = r$.

Simplices are characterized by the property that their points are unique convex combinations of the vertices.

Theorem 1.1.6. *A convex set $A \subset \mathbb{R}^n$ is a simplex, if and only if there exist $x_0, \dots, x_k \in A$ such that each $x \in A$ has a* **unique** *representation as a convex combination of $x_0, \dots, x_k$.*

Proof. By definition, A is a simplex, if $A = \mathrm{conv}\,\{x_0, \dots, x_k\}$ with affinely independent $x_0, \dots, x_k \in \mathbb{R}^n$. The assertion therefore follows from Theorem 1.1.2 together with the uniqueness property of affine combinations (with respect to affinely independent points) and the well-known characterizations of affine independence (see also Exercise 11). □

Exercises and problems

1. (a) Show that $A \subset \mathbb{R}^n$ is convex, if and only if $\alpha A + \beta A = (\alpha + \beta)A$ holds, for all $\alpha, \beta \geq 0$.
   (b) Which non-empty sets $A \subset \mathbb{R}^n$ are characterized by $\alpha A + \beta A = (\alpha + \beta)A$, for all $\alpha, \beta \in \mathbb{R}$?
2. Let $A \subset \mathbb{R}^n$ be closed. Show that A is convex, if and only if $A + A = 2A$ holds.
3. A set
   $$R := \{x + \alpha y : \alpha \geq 0\}, \quad x \in \mathbb{R}^n, y \in S^{n-1},$$
   is called a *ray* (starting in x with direction y).
   Let $A \subset \mathbb{R}^n$ be convex and unbounded. Show that A contains a ray.
   Hint: Start with the case of a closed set A. For the general case, Theorem 1.3.2 is useful.
4. For a set $A \subset \mathbb{R}^n$, the *polar* A° is defined as
   $$A^\circ := \{x \in \mathbb{R}^n : \langle x, y \rangle \leq 1 \; \forall y \in A\}.$$
   Show that:

(a) A° is closed, convex and contains 0.

(b) If $A \subset B$, then $A^\circ \supset B^\circ$.

(c) $(A \cup B)^\circ = A^\circ \cap B^\circ$.

(d) If P is a polytope, P° is polyhedral.

5. (a) If $\|\cdot\|' : \mathbb{R}^n \to [0,\infty)$ is a norm, show that the corresponding unit ball $B' := \{x \in \mathbb{R}^n : \|x\|' \le 1\}$ is convex and symmetric (i.e. $B' = -B'$).

(b) Show that

$$\|\cdot\|_1 : \mathbb{R}^n \to [0,\infty), \quad x = (x^{(1)}, \dots, x^{(n)}) \mapsto \sum_{i=1}^n |x^{(i)}|,$$

and

$$\|\cdot\|_\infty : \mathbb{R}^n \to [0,\infty), \quad x = (x^{(1)}, \dots, x^{(n)}) \mapsto \max_{i=1,\dots,n} |x^{(i)}|,$$

are norms. Describe the corresponding unit balls B_1 and B_∞.

(c) Show that for an arbitrary norm $\|\cdot\|' : \mathbb{R}^n \to [0,\infty)$ there are constants $\alpha, \beta, \gamma > 0$ such that

$$\alpha\|\cdot\|_1 \le \beta\|\cdot\|_\infty \le \|\cdot\|' \le \gamma\|\cdot\|_1.$$

Describe these inequalities in terms of the corresponding unit balls B_1, B_∞, B'.

Hint: Show first the last inequality. Then prove that

$$\inf\{\|x\|_\infty : x \in \mathbb{R}^n, \|x\|' = 1\} > 0,$$

and deduce the second inequality from that.

(d) Use (c) to show that all norms on $\mathbb{R}^n$ are equivalent.

6. For a set $A \subset \mathbb{R}^n$ let

$$\ker A := \{x \in A : [x,y] \subset A \text{ for all } y \in A\}$$

be the *kernel* of A. Show that $\ker A$ is convex. Show by an example that $A \subset B$ does not imply $\ker A \subset \ker B$.

7. Let $A \subset \mathbb{R}^n$ be a *locally finite* set (this means that $A \cap B(r)$ is a finite set, for all $r \ge 0$). For each $x \in A$, we define the *Voronoi cell*

$$C(x,A) := \{z \in \mathbb{R}^n : \|z - x\| \le \|z - y\| \ \forall y \in A\},$$

consisting of all points $z \in \mathbb{R}^n$ which have x as their nearest point (or one of their nearest points) in A.

(a) Show that the Voronoi cells $C(x,A), x \in A$, are closed and convex.

(b) If $\operatorname{conv} A = \mathbb{R}^n$, show that the Voronoi cells $C(x,A), x \in A$, are bounded and polyhedral, hence they are convex polytopes.

Hint: Use Exercise 3.

(c) Show by an example that the condition $\operatorname{conv} A = \mathbb{R}^n$ is not necessary for the boundedness of the Voronoi cells $C(x,A), x \in A$.

8. Show that the set $\mathcal{A}$ of all convex subsets of $\mathbb{R}^n$ is a complete lattice with respect to the inclusion order.

   Hint: Define
$$A \wedge B := A \cap B,$$
$$A \vee B := \text{conv}\,(A \cup B),$$
$$\inf \mathcal{M} := \bigcap_{A \in \mathcal{M}} A, \quad \mathcal{M} \subset \mathcal{A},$$
$$\sup \mathcal{M} := \text{conv}\left(\bigcup_{A \in \mathcal{M}} A\right), \quad \mathcal{M} \subset \mathcal{A}.$$

9. Show that, for $A, B \subset \mathbb{R}^n$, we have $\text{conv}\,(A + B) = \text{conv}\, A + \text{conv}\, B$.

10. Let $A, B \subset \mathbb{R}^n$ be nonempty convex sets, and let $x \in \mathbb{R}^n$. Show that

    (a)
$$\text{conv}\,(\{x\} \cup A) = \{\lambda a + (1 - \lambda)x : \lambda \in [0, 1], a \in A\}.$$

    (b) If $A \cap B = \emptyset$, then
$$\text{conv}\,(\{x\} \cup A) \cap B = \emptyset \qquad \text{or} \qquad \text{conv}\,(\{x\} \cup B) \cap A = \emptyset.$$

11. Assume that $x_1, \ldots, x_k \in \mathbb{R}^n$ are such that each $x \in \text{conv}\,\{x_1, \ldots, x_k\}$ is a unique convex combination of $x_1, \ldots, x_k$. Show that $x_1, \ldots, x_k$ are affinely independent.

12. Let $P = \text{conv}\,\{x_0, \ldots, x_n\}$ be an n-simplex in $\mathbb{R}^n$. Denote by E_i the affine hull of $\{x_0, \ldots, x_n\} \setminus \{x_i\}$ and by H_i the closed half-space bounded by E_i and with $x_i \in H_i$, $i = 0, \ldots, n$.

    (a) Show that $x_i \in \text{int}\, H_i$, $i = 0, \ldots, n$.

    (b) Show that $P = \bigcap_{i=0}^{n} H_i$.

    (c) Show that $P \cap E_i$ is an $(n - 1)$-simplex.

1.2 Combinatorial properties

Combinatorial problems arise in connection with polytopes. In the following, however, we discuss problems of general convex sets which are called combinatorial, since they involve the cardinality of points or sets. The most important results in this part of convex geometry (which is called *Combinatorial Geometry*) are the theorems of CARATHÉODORY, HELLY and RADON.

Theorem 1.2.1 (RADON). *Let $x_1, \dots, x_m \in \mathbb{R}^n$ be affinely dependent points. Then there exists a partition $\{1, \dots, m\} = I \cup J$, $I \cap J = \emptyset$, such that*

$$\operatorname{conv}\{x_i : i \in I\} \cap \operatorname{conv}\{x_j : j \in J\} \neq \emptyset.$$

Proof. Let $x_1, \dots, x_m \in \mathbb{R}^n$ be affinely dependent. Then there exist $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, not all zero, such that

$$\sum_{i=1}^{m} \alpha_i x_i = 0 \quad \text{and} \quad \sum_{i=1}^{m} \alpha_i = 0.$$

Define $I := \{i \in \{1, \dots, m\} : \alpha_i \geq 0\}$ and $J := \{1, \dots, m\} \setminus I$. Then

$$\alpha := \sum_{i \in I} \alpha_i = \sum_{j \in J} (-\alpha_j) > 0.$$

Hence

$$y := \sum_{i \in I} \frac{\alpha_i}{\alpha} x_i = \sum_{j \in J} \frac{-\alpha_j}{\alpha} x_j \in \operatorname{conv}\{x_i : i \in I\} \cap \operatorname{conv}\{x_j : j \in J\}.$$

□

Observe that any sequence of $n+2$ points in $\mathbb{R}^n$ is affinely dependent. As a consequence, we next derive HELLY's Theorem (in a particular version). It provides an answer to a question of the following type. Let $A_1, \dots, A_m$ be a sequence of sets such that any s of these sets enjoy a certain property (for instance, having nonempty intersection). Do then all sets of the sequence enjoy this property?

Theorem 1.2.2 (HELLY). *Let $A_1, \dots, A_m$ be convex sets in $\mathbb{R}^n$, $m \geq n+1$. If each $n+1$ of the sets $A_1, \dots, A_m$ have nonempty intersection, then*

$$\bigcap_{i=1}^{m} A_i \neq \emptyset.$$

Proof. We proceed by induction with respect to $m \geq n+1$. For $m = n+1$ there is nothing to show. Let $m \geq n+2$, and assume that the assertion is true for $m-1$ sets. Hence there are

$$x_i \in A_1 \cap \cdots \cap \check{A}_i \cap \cdots \cap A_m$$

(A_i is omitted) for $i = 1, \ldots, m$. The sequence $x_1, \ldots, x_m$ of $m \geq n+2$ points is affinely dependent. By Radon's theorem (possibly after a change of notation) there is some $k \in \{1, \ldots, m-1\}$ and a point $x \in \mathbb{R}^n$ satisfying

$$x \in \operatorname{conv}\{x_1, \ldots, x_k\} \cap \operatorname{conv}\{x_{k+1}, \ldots, x_m\}.$$

Since $x_1, \ldots, x_k \in A_{k+1}, \ldots, A_m$, we get

$$x \in \operatorname{conv}\{x_1, \ldots, x_k\} \subset A_{k+1} \cap \cdots \cap A_m. \tag{2.2}$$

Furthermore, since $x_{k+1}, \ldots, x_m \in A_1, \ldots, A_k$, we also have

$$x \in \operatorname{conv}\{x_{k+1}, \ldots, x_m\} \subset A_1 \cap \cdots \cap A_k. \tag{2.3}$$

Thus (2.2) and (2.3) yield $x \in A_1 \cap \cdots \cap A_m$. □

HELLY's Theorem has interesting applications. For some of them, we refer to the exercises. In general, the theorem cannot be extended to infinite families of convex sets (see Exercise 1). An exception is the case of compact sets.

Theorem 1.2.3 (HELLY). *Let $\mathcal{A}$ be a family of at least $n+1$ compact convex sets in $\mathbb{R}^n$ ($\mathcal{A}$ may be infinite) and assume that any $n+1$ sets in $\mathcal{A}$ have a non-empty intersection. Then, there is a point $x \in \mathbb{R}^n$ which is contained in all sets of $\mathcal{A}$.*

Proof. By Theorem 1.2.2, every finite subfamily of $\mathcal{A}$ has a non-empty intersection. For compact sets, this implies

$$\bigcap_{A \in \mathcal{A}} A \neq \emptyset.$$

In fact, if $\bigcap_{A \in \mathcal{A}} A = \emptyset$, then

$$\bigcup_{A \in \mathcal{A}} (\mathbb{R}^n \setminus A) = \mathbb{R}^n.$$

By the covering property, any compact $A_0 \in \mathcal{A}$ is covered by finitely many open sets $\mathbb{R}^n \setminus A_1, ..., \mathbb{R}^n \setminus A_k$, $A_i \in \mathcal{A}$. This implies

$$\bigcap_{i=0}^{k} A_i = \emptyset,$$

a contradiction. □

The following result will be frequently used later on.

Theorem 1.2.4 (CARATHÉODORY). *For a set $A \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$ the following two assertions are equivalent:*
(a) $x \in \operatorname{conv} A$,
(b) *there is an r-simplex P $(0 \leq r \leq n)$ with vertices in A and such that $x \in P$.*

Proof. (b)⇒(a): Since $\operatorname{vert} P \subset A$, we have $x \in P = \operatorname{conv}\operatorname{vert} P \subset \operatorname{conv} A$.

(a)⇒(b): By Theorem 1.1.2, $x = \alpha_1 x_1 + \cdots + \alpha_k x_k$ with $k \in \mathbb{N}, x_1, \ldots, x_k \in A$, $\alpha_1, \ldots, \alpha_k \in (0, 1]$ and $\alpha_1 + \ldots + \alpha_k = 1$. Let k be the minimal number for which such a representation is possible, i.e. x is not in the convex hull of any $k-1$ points of A. We now show that $x_1, \ldots, x_k$ are affinely independent. In fact, assume that there were numbers $\beta_1, \ldots, \beta_k \in \mathbb{R}$, not all zero, such that

$$\sum_{i=1}^{k} \beta_i x_i = 0 \quad \text{and} \quad \sum_{i=1}^{k} \beta_i = 0.$$

Let J be the set of indices $i \in \{1, \ldots, k\}$, for which $\beta_i > 0$ and choose $i_0 \in J$ such that

$$\frac{\alpha_{i_0}}{\beta_{i_0}} = \min_{i \in J} \frac{\alpha_i}{\beta_i}.$$

Then, we have

$$x = \sum_{i=1}^{k} \left(\alpha_i - \frac{\alpha_{i_0}}{\beta_{i_0}} \beta_i \right) x_i$$

with

$$\alpha_i - \frac{\alpha_{i_0}}{\beta_{i_0}} \beta_i \geq 0, \quad \sum_{i=1}^{k} \left(\alpha_i - \frac{\alpha_{i_0}}{\beta_{i_0}} \beta_i \right) = 1 \quad \text{and} \quad \alpha_{i_0} - \frac{\alpha_{i_0}}{\beta_{i_0}} \beta_{i_0} = 0.$$

This is a contradiction to the minimality of k. □

Exercises and problems

1. Show by an example that Theorem 1.2.3 is wrong if the sets in $\mathcal{A}$ are only assumed to be closed (and not necessarily compact).

2. In an old German fairy tale, a tailor claimed the fame to have 'killed seven with one stroke'. A closer examination showed that the victims were in fact flies which had landed on a toast covered with jam. The tailor had used a fly-catcher of convex shape for his sensational victory. As the remains of the flies on the toast showed, it was possible to kill any three of them with one stroke of the (suitably) shifted fly-catcher without even turning the direction of the handle.

   Is it possible that the tailor told the truth?

3. Let $\mathcal{F}$ be a family of finitely many parallel closed segments in $\mathbb{R}^2$, $|\mathcal{F}| \geq 3$. Suppose that for any three segments in $\mathcal{F}$ there is a line intersecting all three segments.

   Show that there is a line in $\mathbb{R}^2$ intersecting all the segments in $\mathcal{F}$.

$*$ Show that the above result remains true without the finiteness condition.

4. Prove the following version of CARATHÉODORY's theorem:

   Let $A \subset \mathbb{R}^n$ and $x_0 \in A$ be fixed. Then $\operatorname{conv} A$ is the union of all simplices with vertices in A and such that x_0 is one of the vertices.

$*$ 5. Prove the following generalization of CARATHÉODORY's theorem (Theorem of BUNDT):

   Let $A \subset \mathbb{R}^n$ be a connected set. Then $\operatorname{conv} A$ is the union of all simplices with vertices in A and dimension at most $n - 1$.

6. Collect further examples for applications of Helly's theorem:

   Lutwak's containment result (simplices),

   centre point result

   elementary applications

1.3 Topological properties

Although convexity is a purely algebraic property, it has a variety of topological consequences. One striking property of convex sets is that they always have (relative) interior points. In order to prove that, we first need an auxiliary result.

Proposition 1.3.1. *If* $P = \mathrm{conv}\,\{x_0, \dots, x_k\}$ *is a* k*-simplex in* $\mathbb{R}^n$*,* $1 \leq k \leq n$*, then*

$$\mathrm{rel\,int}\, P = \{\alpha_0 x_0 + \dots + \alpha_k x_k : \alpha_i \in (0,1), \alpha_0 + \ldots + \alpha_k = 1\}\,.$$

Proof. W.l.o.g. we may assume $k = n$ (working in aff A) and $x_0 = 0$ (using $\alpha_0 = 1 - \alpha_1 - \ldots - \alpha_k$ and replacing P by $P - x_0$). Then we have

$$P = \{\alpha_1 x_1 + \dots + \alpha_n x_n : \alpha_i \in [0,1], \alpha_1 + \ldots + \alpha_k \leq 1\}\,,$$

and we need to show that

$$\mathrm{int}\, P = \{\alpha_1 x_1 + \dots + \alpha_n x_n : \alpha_i \in (0,1), \alpha_1 + \ldots + \alpha_k < 1\}\,.$$

Notice that $x_1, \dots, x_n$ is a basis of $\mathbb{R}^n$. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $F(x) = (\alpha_1, \dots, \alpha_k)$ if $x = \alpha_1 x_1 + \ldots + \alpha_k x_k$. Then F is a homeomorphism. Therefore, $\mathrm{int}\, P = F^{-1}(\mathrm{int}\, F(P))$. Obviously,

$$\mathrm{int}\, F(P) = \{(\alpha_1, ..., \alpha_n) : \alpha_i \in (0,1), \alpha_1 + \ldots + \alpha_k < 1\}$$

and the proof is complete. □

Theorem 1.3.2. *If* $A \subset \mathbb{R}^n$*,* $A \neq \emptyset$*, is convex, then* $\mathrm{rel\,int}\, A \neq \emptyset$*.*

Proof. If $\dim A = k$, then A contains $k+1$ affinely independent points and hence a k-simplex P. By Proposition 1.3.1, there is some $x \in \mathrm{rel\,int}\, P$. Each such x fulfills $x \in \mathrm{rel\,int}\, A$. □

Theorem 1.3.2 shows that, for the investigation of a fixed convex set A, it is useful to consider the affine hull of A, as the basic space, since then A has interior points. We will often take advantage of this fact by assuming that the affine hull of A is the whole space $\mathbb{R}^n$. Therefore, proofs in the following frequently start with the sentence that we may assume (w.l.o.g.) that the convex set under consideration has dimension n.

A further consequence of convexity is that topological notions like interior or closure of a (convex) set can be expressed in purely geometric terms.

Theorem 1.3.3. *If* $A \subset \mathbb{R}^n$ *is convex, then*

$$\mathrm{cl}\, A = \{x \in \mathbb{R}^n : \exists y \in A \text{ with } [y,x) \subset A\}$$

and

$$\mathrm{int}\, A = \{x \in \mathbb{R}^n : \forall y \in \mathbb{R}^n \setminus \{x\}\, \exists z \in (x,y) \text{ with } [x,z] \subset A\}.$$

Again, we first need an auxiliary result.

Proposition 1.3.4. *If* $A \subset \mathbb{R}^n$ *is convex,* $x \in \operatorname{cl} A$, $y \in \operatorname{relint} A$, *then* $[y, x) \subset \operatorname{relint} A$.

Proof. As we explained above, we may assume $\dim A = n$. Let $x \in \operatorname{cl} A, y \in \operatorname{relint} A$ and $z \in (y, x)$, that is $z = \alpha y + (1-\alpha)x$, $\alpha \in (0, 1)$. We have to show that $z \in \operatorname{int} A$. Since $x \in \operatorname{cl} A$, there exists a sequence $x_k \to x$ with $x_k \in A$ for $k \in \mathbb{N}$. Then $y_k := \frac{1}{\alpha}(z - (1-\alpha)x_k)$ converges towards y, as $k \to \infty$. Since $y \in \operatorname{int} A$, for k large enough we have $y_k \in \operatorname{int} A$. Then, there exists an open ball V around y_k with $V \subset A$. The convexity of A implies $z \in \alpha V + (1-\alpha)x_k \subset A$. Since $\alpha V + (1-\alpha)x_k$ is open, $z \in \operatorname{int} A$. □

Proof of Theorem 1.3.3. The case $A = \emptyset$ is trivial, hence we assume now that $A \neq \emptyset$.

Concerning the first equation, we may assume $\dim A = n$ since the sets on both sides depend only on $\operatorname{aff} A$. Let B be the set on the right-hand side. Then we obviously have $B \subset \operatorname{cl} A$. To show the converse inclusion, let $x \in \operatorname{cl} A$. By Theorem 1.3.2 there is a point $y \in \operatorname{int} A$, hence by Proposition 1.3.4 we have $[y, x) \subset \operatorname{int} A \subset A$. Therefore, $x \in B$.

The second equation is trivial for $\dim A < n$, since then both sides are empty. Hence, let $\dim A = n$. We denote the set on the right-hand side by C. Then the inclusion $\operatorname{int} A \subset C$ is obvious. For the converse, let $x \in C$. Again, we choose $y \in \operatorname{int} A$ by Theorem 1.3.2, $y \neq x$. The definition of C implies that for $2x - y \in \mathbb{R}^n$ there exists $z \in (x, 2x - y)$ with $z \in A$. Then $x \in (y, z)$ and Proposition 1.3.4 shows that $x \in \operatorname{int} A$. □

Remarks. (1) For simplicity, we have formulated Theorem 1.3.3(b) for the interior of a convex set A. The result can be easily modified to cover the case of the relative interior of a lower dimensional set A.

(2) Theorem 1.3.3 shows that (and how) topological notions like the interior and the closure of a set can be defined for convex sets A on a purely algebraic basis, without that a topology has to be given in the underlying space. This can be used in arbitrary real vector spaces V (without a given topology) to introduce and study topological properties of convex sets.

In view of this remark, we deduce the following two corollaries from Theorem 1.3.3, instead of giving a direct proof based on the topological notions relint and cl.

Corollary 1.3.5. *For a convex set* $A \subset \mathbb{R}^n$, *the sets* $\operatorname{relint} A$ *and* $\operatorname{cl} A$ *are convex.*

Proof. The convexity of $\operatorname{relint} A$ follows immediately from Proposition 1.3.4.

For the convexity of $\operatorname{cl} A$, let $A \neq \emptyset$, $x_1, x_2 \in \operatorname{cl} A$, $\alpha \in (0, 1)$. From Theorem 1.3.3, we get points $y_1, y_2 \in A$ with $[y_1, x_1) \subset A$, $[y_2, x_2) \subset A$. Hence

$$\alpha[y_1, x_1) + (1-\alpha)[y_2, x_2) \subset A.$$

Since

$$[\alpha y_1 + (1-\alpha)y_2, \alpha x_1 + (1-\alpha)x_2) \subset \alpha[y_1, x_1) + (1-\alpha)[y_2, x_2),$$

we obtain $\alpha x_1 + (1-\alpha)x_2 \in \operatorname{cl} A$, again from Theorem 1.3.3. □

Corollary 1.3.6. *For a convex set $A \subset \mathbb{R}^n$,*

$$\operatorname{cl} A = \operatorname{cl} \operatorname{rel\,int} A$$

and

$$\operatorname{rel\,int} A = \operatorname{rel\,int} \operatorname{cl} A.$$

Proof. The inclusion

$$\operatorname{cl} \operatorname{rel\,int} A \subset \operatorname{cl} A$$

is obvious. Let $x \in \operatorname{cl} A$. By Theorem 1.3.2 there is a $y \in \operatorname{rel\,int} A$ and by Proposition 1.3.4 we have $[y, x) \subset \operatorname{rel\,int} A$. Since $\operatorname{rel\,int} A$ is convex (Corollary 1.3.5), Theorem 1.3.3 implies $x \in \operatorname{cl} \operatorname{rel\,int} A$.

The inclusion

$$\operatorname{rel\,int} A \subset \operatorname{rel\,int} \operatorname{cl} A$$

is again obvious. Let $x \in \operatorname{rel\,int} \operatorname{cl} A$. Since $\operatorname{cl} A$ is convex (Corollary 1.3.5), we can apply Theorem 1.3.3 in aff $A =$ aff cl A to $\operatorname{cl} A$. Therefore, for $y \in \operatorname{rel\,int} A$ (which exists by Theorem 1.3.2), $y \neq x$, we obtain $z \in \operatorname{cl} A$ such that $x \in (y, z)$. By Proposition 1.3.4, $x \in \operatorname{rel\,int} A$. □

We finally study the topological properties of the convex hull operator. For a closed set $A \subset \mathbb{R}^n$, the convex hull $\operatorname{conv} A$ need not be closed. A simple example is given by the set

$$A := \{(t, t^{-1}) : t > 0\} \cup \{(0, 0)\} \subset \mathbb{R}^2.$$

However, the convex hull operator behaves well with respect to open and compact sets.

Theorem 1.3.7. *If $A \subset \mathbb{R}^n$ is open, $\operatorname{conv} A$ is open. If $A \subset \mathbb{R}^n$ is compact, $\operatorname{conv} A$ is compact.*

Proof. Let A be open and $x \in \operatorname{conv} A$. Then there exist $x_i \in A$ and $\alpha_i \in (0, 1]$, $i \in \{1, \dots, k\}$, such that $x = \alpha_1 x_1 + \cdots + \alpha_k x_k$ and $\alpha_1 + \ldots + \alpha_k = 1$. We can choose a ball U around the origin such that $x_i + U \subset A \subset \operatorname{conv} A$, $i = 1, \dots, k$. Since

$$U + x = \alpha_1(U + x_1) + \cdots + \alpha_k(U + x_k) \subset \operatorname{conv} A,$$

we have $x \in \operatorname{int} \operatorname{conv} A$, hence $\operatorname{conv} A$ is open.

Now let A be compact. Since A is contained in a ball $B(r)$, we have $\operatorname{conv} A \subset B(r)$, i.e. $\operatorname{conv} A$ is bounded. In order to show that $\operatorname{conv} A$ is closed, let $x_k \to x$, $x_k \in \operatorname{conv} A$, $k \in \mathbb{N}$. By Theorem 1.2.4, each x_k has a representation

$$x_k = \alpha_{k0} x_{k0} + \cdots + \alpha_{kn} x_{kn}$$

with

$$\alpha_{ki} \in [0, 1], \quad \sum_{i=0}^{n} \alpha_{ki} = 1 \quad \text{and} \quad x_{ki} \in A.$$

Because A and $[0, 1]$ are compact, we find a subsequence $(k_r)_{r \in \mathbb{N}}$ in $\mathbb{N}$ such that the $2n + 2$ sequences $(x_{k_r j})_{r \in \mathbb{N}}, j = 0, \dots, n$, and $(\alpha_{k_r j})_{r \in \mathbb{N}}, j = 0, \dots, n$, all converge. We denote the limits by y_j and β_j, $j = 0, \dots, n$. Then, $y_j \in A$, $\beta_j \in [0, 1]$, $\beta_0 + \ldots + \beta_n = 1$ and $x = \beta_0 y_0 + \cdots \beta_n y_n$. Hence, $x \in \operatorname{conv} A$. □

Remarks. **(1)** The last theorem shows, in particular, that a convex polytope P is compact; a fact, which can of course be proved in a simpler, more direct way.

(2) We give an alternative argument for the first part of Theorem 1.3.7 (following a suggestion of Mathew Penrose). Let A be open and $x \in \operatorname{conv} A$. Then there exist $x_i \in A$ and $\alpha_i \in (0, 1]$, $i \in \{1, \dots, k\}$, such that $x = \alpha_1 x_1 + \cdots + \alpha_k x_k$ and $\alpha_1 + \ldots + \alpha_k = 1$. If $k = 1$, the assertion is clear. If $k \geq 2$, we have

$$x = \alpha_1 x_1 + (1 - \alpha_1) \underbrace{\sum_{j=2}^{k} \frac{\alpha_j}{1 - \alpha_1} x_j}_{=:y} .$$

Since $x_1 \in \operatorname{int} \operatorname{conv} A$ amd $y \in \operatorname{conv} A$, Proposition 1.3.4 yields that $x \in [x_1, y) \subset \operatorname{int} \operatorname{conv} A$.

(3) For an alternative argument for the second part of Theorem 1.3.7, define

$$C := \{(\alpha_0, \dots, \alpha_n, x_0, \dots, x_n) \in [0, 1]^{n+1} \times A^{n+1} : \alpha_0 + \ldots + \alpha_n = 1\}$$

and

$$f : C \to \operatorname{conv} A, \quad f(\alpha_0, \dots, \alpha_n, x_0, \dots, x_n) := \sum_{i=0}^{n} \alpha_i x_i.$$

Clearly, f is continuous and C is compact. Hence $f(C)$ is compact. By Carathéodory's theorem, $f(C) = \operatorname{conv} A$, which shows that $\operatorname{conv} A$ is compact.

Exercises and problems

1. Let $P = \operatorname{conv} \{a_0, \dots, a_n\}$ be an n-simplex in $\mathbb{R}^n$ and $x \in \operatorname{int} P$.

   Show that the polytopes

   $$P_i := \operatorname{conv} \{a_0, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n\}, \quad i = 0, \dots, n,$$

   are n-simplices with pairwise disjoint interiors and that

   $$P = \bigcup_{i=0}^{n} P_i.$$

2. Show that, for $A \subset \mathbb{R}^n$,

   $$\operatorname{cl} \operatorname{conv} A = \bigcap \{B \subset \mathbb{R}^n : B \supset A,\ B \text{ closed and convex}\}.$$

3. Let $A, B \subset \mathbb{R}^n$ be convex.

   (a) Show that $\operatorname{relint} (A + B) = \operatorname{relint} A + \operatorname{relint} B$.

   (b) If A (or B) is bounded, show that $\operatorname{cl} (A + B) = \operatorname{cl} A + \operatorname{cl} B$.

   (c) Show by an example that (b) is wrong, if neither A nor B are assumed to be bounded.

4. Let $A, B \subset \mathbb{R}^n$ be convex, A closed, B compact. Show that $A + B$ is closed (and convex). Give an example which shows that the compactness of one of the sets A, B is necessary for this statement.

1.4 Support and separation theorems

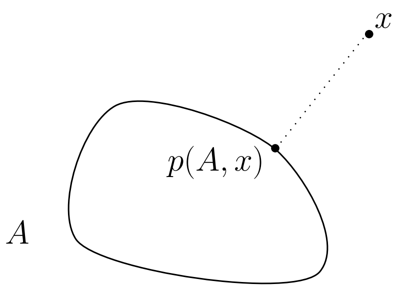
Convex sets are sets which contain with their elements also all convex combinations. In this section, we consider a description of convex sets which is of a dual nature, in that it describes convex sets A as intersections of half-spaces. For such a result, we have to assume that A is a closed set.

We start with results on the metric projection which are of independent interest.

Theorem 1.4.1. *Let $A \subset \mathbb{R}^n$ be nonempty, convex and closed. Then for each $x \in \mathbb{R}^n$, there is a unique point $p(A, x) \in A$ satisfying*

$$\|p(A, x) - x\| = \inf_{y \in A} \|y - x\|.$$

Definition. The mapping $p(A, \cdot) : \mathbb{R}^n \to A$ is called the *metric projection* (onto A).



Proof of Theorem 1.4.1. For $x \in A$, we obviously have $p(A, x) = x$. For $x \notin A$, there is a ball $B(r)$ such that

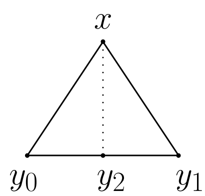
$$A \cap (x + B(r)) \neq \emptyset.$$

Then,

$$\inf_{y \in A} \|y - x\| = \inf_{y \in A \cap (x + B(r))} \|y - x\|.$$

Since $A_r := A \cap (x + B(r))$ is compact and $f : y \mapsto \|y - x\|$ continuous, there is a point $y_0 \in A$ realizing the minimum of f on A_r.

If $y_1 \in A$ is a second point realizing this minimum, with $y_1 \neq y_0$, then $y_2 := \frac{1}{2}(y_0 + y_1) \in A$ and $\|y_2 - x\| < \|y_0 - x\|$, by Pythagoras' theorem.



This is a contradiction and hence the metric projection $p(A, x)$ is unique. $\square$

Remark. As the above proof shows, the existence of a nearest point $p(A,x)$ is guaranteed for all closed sets A. The convexity of A is responsible for the uniqueness of $p(A,x)$. A more general class of sets consists of closed sets A, for which the uniqueness of $p(A,x)$ holds at least in an ε-neighborhood of A, i.e. for all $x \in A + \varepsilon B^n$, with $\varepsilon > 0$. Such sets are called sets of *positive reach*, and the largest ε for which uniqueness of the metric projection holds is called the *reach* of A. Convex sets thus have reach ∞.

Definition. Let $A \subset \mathbb{R}^n$ be closed and convex, and let $E = \{f = \alpha\}$ be a hyperplane. E is called *supporting hyperplane* of A, if $A \cap E \neq \emptyset$ and A is contained in one of the two closed half-spaces $\{f \leq \alpha\}$, $\{f \geq \alpha\}$ (or in both, but this implies $A \subset \{f = \alpha\}$, hence it is only possible for lower dimensional sets A). A half-space containing A and bounded by a supporting hyperplane of A is called *supporting half-space* of A, the set $A \cap E$ is called *support set* and any $x \in A \cap E$ is called *supporting point*.

If E is a supporting hyperplane of A, we also say shortly that the hyperplane E supports A.

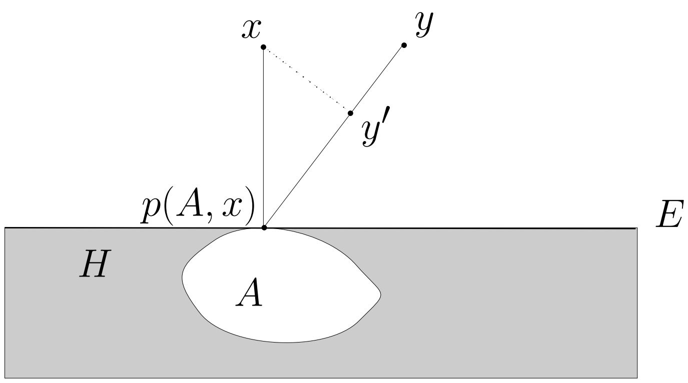
Example. The set

$$A := \{(x^{(1)}, x^{(2)}) \in \mathbb{R}^2 : x^{(2)} \geq \frac{1}{x^{(1)}},\, x^{(1)} > 0\}$$

is closed and convex. The line $g := \{x^{(1)} + x^{(2)} = 2\}$ is a supporting line, since $(1,1) \in A \cap g$ and $A \subset \{x^{(1)} + x^{(2)} \geq 2\}$. The lines $h := \{x^{(1)} = 0\}$ and $k := \{x^{(2)} = 0\}$ bound the set A, but are not supporting lines since they do not have a point in common with A.

Theorem 1.4.2. *Let $A \subset \mathbb{R}^n$ be nonempty, closed and convex and let $x \in \mathbb{R}^n \setminus A$. Then, the hyperplane E through $p(A,x)$, orthogonal to $x - p(A,x)$, supports A. Moreover, the half-space H bounded by E and not containing x is a supporting half-space.*

Proof. Obviously $x \notin E$. Since $p(A,x) \in E \cap A$, it remains to show that $A \subset H$. Assume that there is $y \in A, y \notin H$. Then $\langle y - p(A,x), x - p(A,x)\rangle > 0$. We consider the orthogonal projection $\bar{y}$ of x onto the line through $p(A,x)$ and y. By Pythagoras' theorem, $\|\bar{y} - x\| < \|p(A,x) - x\|$. If $\bar{y} \in (p(A,x), y]$, we put $y' := \bar{y}$. Otherwise, we have $y \in (p(A,x), \bar{y}]$ and put $y' := y$.



In both cases we obtain a point $y' \in (p(A,x), y] \subset A$ with $\|y' - x\| < \|p(A,x) - x\|$. This is a contradiction, hence we conclude $A \subset H$. □

Corollary 1.4.3. *Every nonempty, closed convex set $A \subset \mathbb{R}^n$, $A \neq \mathbb{R}^n$, is the intersection of all closed half-spaces which contain A. More specifically, A is the intersection of all its supporting half-spaces.*

Proof. Obviously, A lies in the intersection B of its supporting half-spaces. For $x \notin A$, Theorem 1.4.2 implies the existence of a supporting half-space H of A with $x \notin H$. Hence $x \notin B$. □

Theorem 1.4.2 and Corollary 1.4.3 do not imply that every boundary point of A is a support point. In order to show such a result, we approximate $x \in \operatorname{bd} A$ by points x_k from $\mathbb{R}^n \setminus A$ and consider the corresponding supporting hyperplanes E_k which exist by Theorem 1.4.2. For $x_k \to x$, we want to define a supporting hyperplane in x as the limit of the E_k. A first step in this direction is to show that $p(A, x_k) \to p(A, x)$ (where $p(A, x) = x$), hence to show that $p(A, \cdot)$ is continuous. We even show now that $p(A, \cdot)$ is Lipschitz continuous with Lipschitz constant 1.

Theorem 1.4.4. *Let $A \subset \mathbb{R}^n$ be nonempty, closed and convex. Then,*

$$\|p(A, x) - p(A, y)\| \leq \|x - y\|,$$

for all $x, y \in \mathbb{R}^n$.

Proof. During the proof, we abbreviate $p(A, \cdot)$ by p. Let $x, y \in \mathbb{R}^n$. The case $x \in A$ or $y \in A$ is easy, thus we assume now $x, y \notin A$. Then, by Theorem 1.4.2, we obtain $\langle x - p(x), p(y) - p(x)\rangle \leq 0$ and $\langle y - p(y), p(x) - p(y)\rangle \leq 0$. Addition of these two inequalities yields

$$\langle p(y) - p(x), p(y) - y + x - p(x)\rangle \leq 0,$$

and therefore

$$\|p(y) - p(x)\|^2 \leq \langle p(y) - p(x), y - x\rangle \leq \|p(y) - p(x)\| \cdot \|y - x\|,$$

where the Cauchy-Schwarz inequality was used for the last estimate. For $p(x) \neq p(y)$, this implies the required inequality. The case $p(x) = p(y)$ is trivial. □

Theorem 1.4.5 (Support Theorem)**.** *Let $A \subset \mathbb{R}^n$ be closed and convex. Then through each boundary point of A there exists a supporting hyperplane.*

Proof. For given $x \in \operatorname{bd} A$, we consider the closed unit ball $x + B(1)$ around x. For each $k \in \mathbb{N}$, we choose $x_k \in x + B(1)$, $x_k \notin A$, and such that $\|x - x_k\| < \frac{1}{k}$. Then

$$\|x - p(A, x_k)\| = \|p(A, x) - p(A, x_k)\| \leq \|x - x_k\| < \frac{1}{k},$$

by Theorem 1.4.4. Since $x_k, p(A, x_k)$ are interior points of $x + B(1)$, there is a (unique) boundary point y_k in $x + B(1)$ such that $x_k \in (p(A, x_k), y_k)$. Theorem 1.4.2 then implies $p(A, y_k) = p(A, x_k)$. In view of the compactness of $x + B(1)$, we may choose a converging subsequence $y_{k_r} \to y$. By Theorem 1.4.4, $p(A, y_{k_r}) \to p(A, y)$ and $p(A, y_{k_r}) = p(A, x_{k_r}) \to p(A, x) = x$, hence $p(A, y) = x$. Since $y \in \operatorname{bd}(x + B(1))$, we also know that $x \neq y$. The assertion now follows from Theorem 1.4.2. □

Remark. Supporting hyperplanes, half-spaces and points can be defined for nonconvex sets A as well; they only exist however, if $\operatorname{conv} A$ is closed and not all of $\mathbb{R}^n$. Then, $\operatorname{conv} A$ is the intersection of all supporting half-spaces of A.

Some of the previous results can be interpreted as separation theorems. For two sets $A, B \subset \mathbb{R}^n$ and a hyperplane $E = \{f = \alpha\}$, we say that E *separates* A and B, if either $A \subset \{f \leq \alpha\}, B \subset \{f \geq \alpha\}$ or $A \subset \{f \geq \alpha\}, B \subset \{f \leq \alpha\}$. Theorem 1.4.2 then says that a closed convex set A and a point $x \notin A$ can be separated by a hyperplane (there is even a separating hyperplane which has positive distance to both, A and x). This result can be extended to compact convex sets B (instead of the point x). Theorem 1.4.5 says that each boundary point of A can be separated from A by a hyperplane. The following result gives a general criterion for sets, which can be separated.

Theorem 1.4.6 (Separation Theorem)**.** *Let $A, B \subset \mathbb{R}^n$ be nonempty and convex with*

$$\operatorname{rel\,int} A \cap \operatorname{rel\,int} B = \emptyset.$$

Then, there exists a hyperplane E which separates A and B.

Proof. Assume $0 \in \operatorname{rel\,int} A - \operatorname{rel\,int} B$. Then, there is a point $x \in \operatorname{rel\,int} A$ with $-x \in -\operatorname{rel\,int} B$, hence $x \in \operatorname{rel\,int} B$. Thus, $x \in \operatorname{rel\,int} A \cap \operatorname{rel\,int} B$, a contradiction. It follows that $0 \notin \operatorname{rel\,int} A - \operatorname{rel\,int} B = \operatorname{rel\,int}(A - B)$ (see Exercise 1.3.3(a)).

If $0 \notin \operatorname{cl}(A - B)$, we apply Theorem 1.4.2 (in $\operatorname{aff}(A - B)$). If $0 \in \operatorname{cl}(A - B)$, we apply Theorem 1.4.5 (in $\operatorname{aff}(A - B)$). In both cases, we obtain a hyperplane $E = \{f = 0\}$ through 0 with $A - B \subset \{f \leq 0\}$. Put $\alpha := \sup_{x \in A} f(x)$, then $A \subset \{f \leq \alpha\}$. Let $y \in B$. Then, for any $x \in A$, $f(x) - f(y) = f(x - y) \leq 0$ and thus $f(y) \geq f(x)$ for all $x \in A$. This shows that $f(y) \geq \alpha$, i.e. $B \subset \{f \geq \alpha\}$. □

Remarks. (1) In topological vector spaces V of infinite dimensions similar support and separation theorems hold true, however there are some important differences, mainly due to the fact that convex sets A in V need not have relative interior points. Therefore a common assumption is that $\operatorname{int} A \neq \emptyset$. Otherwise it is possible that A is closed but does not have any support points, or, in the other direction, that every point of A is a support point (although A does not lie in a hyperplane).

(2) Some of the properties which we derived are characteristic for convexity. For example, a closed set $A \subset \mathbb{R}^n$ such that each $x \notin A$ has a unique metric projection onto A, must be convex (Motzkin's Theorem). Also the Support Theorem has a converse. A closed set $A \subset \mathbb{R}^n$, $\operatorname{int} A \neq \emptyset$, such that each boundary point is a support point, must also be convex. For proofs of these results, see e.g. [S, Theorem 1.2.4] or [We].

(3) Let $A \subset \mathbb{R}^n$ be nonempty, closed and convex. Then, for each direction $u \in S^{n-1}$, there is a supporting hyperplane $E(u)$ of A in direction u (i.e. with outer normal u), if and only if A is compact.

For the rest of this section, we consider convex polytopes and show that for a polytope P finitely many supporting half-spaces suffice to generate P (as the intersection). In other words, we show that polytopes are polyhedral sets. First, we introduce the faces of a polytope.

Definition. The support sets of a polytope P are called *faces*. A face F of P is called a k-*face*, if $\dim F = k$, $k \in \{0, \dots, n-1\}$.

Theorem 1.4.7. *The* 0*-faces of a polytope* $P \subset \mathbb{R}^n$ *are given by the vertices of* P*, i.e. they are of the form* $\{x\}$*,* $x \in \operatorname{vert} P$*.*

Proof. Let $\{x\}$ be a 0-face of P. Hence there is a supporting hyperplane $\{f = \alpha\}$ such that $P \subset \{f \leq \alpha\}$ and $P \cap \{f = \alpha\} = \{x\}$. Then $P \setminus \{x\} = P \cap \{f < \alpha\}$ is convex, hence $x \in \operatorname{vert} P$.

Conversely, let $x \in \operatorname{vert} P$ and let $\operatorname{vert} P \setminus \{x\} = \{x_1, \dots, x_k\}$. Then, $x \notin P' := \operatorname{conv}\{x_1, \dots, x_k\}$. By Theorem 1.4.2 there exists a supporting hyperplane $\{f = \alpha\}$ of P' through $p(P', x)$ with supporting half-space $\{f \leq \alpha\}$ and such that $\beta := f(x) > \alpha$. Let $y \in P$, i.e.

$$y = \sum_{i=1}^{k} \alpha_i x_i + \alpha_{k+1} x, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{k+1} \alpha_i = 1.$$

Then

$$f(y) = \sum_{i=1}^{k} \alpha_i \underbrace{f(x_i)}_{\leq \alpha < \beta} + \alpha_{k+1} f(x) \leq \beta$$

and equality holds if and only if $\alpha_1 = \ldots = \alpha_k = 0$ and $\alpha_{k+1} = 1$, i.e. $y = x$. Hence $\{f \leq \beta\}$ is a supporting halfspace and $P \cap \{f = \beta\} = \{x\}$, thus x is a 0-face of P. □

Definition. The 1-faces of a polytope are called *edges*, and the $(n-1)$-faces are called *facets*.

Remark. In the following, we shall not distinguish between 0-faces and vertices anymore, although one is a set and the other is a point.

Theorem 1.4.8. *Let* $P \subset \mathbb{R}^n$ *be a polytope with* $\operatorname{vert} P = \{x_1, \dots, x_k\}$ *and let* F *be a face of* P*. Then,* $F = \operatorname{conv}\{x_i : x_i \in F\}$*.*

Proof. Assume $F = P \cap \{f = \alpha\}$ and, w.l.o.g., $x_1, \dots x_m \in F$ and $x_{m+1}, \dots, x_k \notin F$. If $\{f \leq \alpha\}$ is the supporting half-space, we have $x_{m+1}, \dots, x_k \in \{f < \alpha\}$, i.e. $f(x_j) = \alpha - \delta_j$, $\delta_j > 0$, $j = m+1, \dots, k$.

Let $x \in P$, $x = \alpha_1 x_1 + \cdots + \alpha_k x_k$, $\alpha_i \geq 0$, $\sum \alpha_i = 1$. Then,

$$f(x) = \alpha_1 f(x_1) + \cdots + \alpha_k f(x_k) = \alpha - \alpha_{m+1}\delta_{m+1} - \cdots - \alpha_k \delta_k.$$

Hence, $x \in F$, if and only if $\alpha_{m+1} = \cdots = \alpha_k = 0$. □

Remark. Theorem 1.4.8 implies, in particular, that a face of a polytope is a polytope and that there are only finitely many faces.

Corollary 1.4.9. *A polytope* P *is polyhedral.*

Proof. If $\dim P = k < n$, we can assume w.l.o.g. that $0 \in E := \operatorname{aff} P$. Also, it is possible to write E as an intersection of half-spaces $\tilde{H}_1, \dots, \tilde{H}_j$ in $\mathbb{R}^n$, $E = \bigcap_{j=1}^r \tilde{H}_j$. If P is polyhedral in E, i.e.

$$P = \bigcap_{i=1}^m H_i,$$

where $H_i \subset E$ are k-dimensional half-spaces, then

$$P = \bigcap_{i=1}^m (H_i \oplus E^\perp) \cap \bigcap_{j=1}^r \tilde{H}_j,$$

hence P is polyhedral in $\mathbb{R}^n$. Therefore, it is sufficient to treat the case $\dim P = n$.

Let $F_1, \dots, F_m$ be the faces of P and $H_1, \dots, H_m$ corresponding supporting half-spaces (i.e. half-spaces with $P \subset H_i$ and $F_i = P \cap \operatorname{bd} H_i$, $i = 1, \dots, m$). Then we have

$$P \subset H_1 \cap \dots \cap H_m =: P'.$$

Assume, there is $x \in P' \setminus P$. We choose $y \in \operatorname{int} P$ and consider $[y, x] \cap P$. Since P is compact and convex (and $x \notin P$), there is $z \in (y, x)$ with $\{z\} = [y, x] \cap \operatorname{bd} P$. By the support theorem there is a supporting hyperplane of P through z, and hence there is a face F_i of P with $z \in F_i$. Since each F_i lies in the boundary of P', we have $z \in \operatorname{bd} P'$. On the other hand, Proposition 1.3.4 shows that $z \in \operatorname{int} P'$, a contradiction. □

Exercises and problems

1. Let $A \subset \mathbb{R}^n$ be closed and $\operatorname{int} A \neq \emptyset$. Show that A is convex, if and only if every boundary point of A is a support point.

∗ 2. Let $A \subset \mathbb{R}^n$ be closed. Suppose that for each $x \in \mathbb{R}^n$ the metric projection $p(A, x)$ onto A is uniquely determined. Show that A is convex (MOTZKIN's theorem).

3. Let $A \subset \mathbb{R}^n$ be non-empty, closed and convex. Show that A is compact, if and only if, for any direction $u \in S^{n-1}$, there is a supporting hyperplane $E(u)$ of A in direction u (i.e. with outer normal u).

4. Let $A, K \subset \mathbb{R}^n$ be convex, A closed, K compact, and assume $A \cap K = \emptyset$.

   Show that there is a hyperplane $\{f = \alpha\}$ with $A \subset \{f < \alpha\}$ and $B \subset \{f > \alpha\}$. Show more generally that α can be chosen such that there is an $\epsilon > 0$ with $A \subset \{f \leq \alpha - \epsilon\}$ and $B \subset \{f \geq \alpha + \epsilon\}$ (*strong separation*).

5. A bavarian farmer is happy owner of a large herd of happy cows, consisting of totally black and totally white animals. One day he finds them sleeping in the sun on his largest meadow. Watching

them, he notices that, for any four cows it would be possible to build a straight fence, separating the black cows from the white ones.

Show that the farmer could build a straight fence, separating the whole herd into black and white animals.

Hint: Cows are lazy. When they sleep, they sleep - even if you build a fence across the meadow.

6. Let $F_1, \ldots, F_m$ be the facets of the polytope P and $H_1, \ldots, H_m$ the corresponding supporting half-spaces. Show that

$$(*) \qquad P = \bigcap_{i=1}^{m} H_i.$$

(This is a generalization of the representation shown in the proof of Corollary 1.4.9.) Show further that the representation $(*)$ is minimal in the sense that, for each representation

$$P = \bigcap_{i \in I} \tilde{H}_i,$$

with a family of half-spaces $\{\tilde{H}_i : i \in I\}$, we have $\{H_1, \ldots, H_m\} \subset \{\tilde{H}_i : i \in I\}$.

1.5 Extremal representations

In the previous section we have seen that the trivial representation of closed convex sets $A \subset \mathbb{R}^n$ as intersection of all closed convex sets containing A can be improved to a nontrivial one, where A is represented as the intersection of the supporting half-spaces. On the other hand, we have the trivial representation of A as the set of all convex combinations of points of A. Therefore, we discuss now the similar nontrivial problem to find a subset $B \subset A$, as small as possible, for which $A = \operatorname{conv} B$ holds. Although there are some general results for closed convex sets A, we will concentrate on the compact case, where we can give a complete (and simple) solution for this problem.

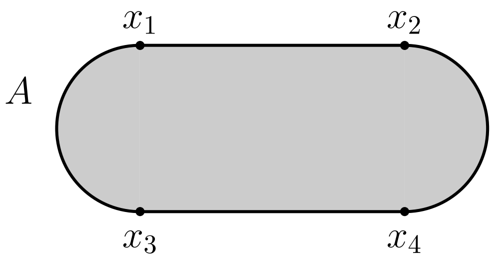
Definition. Let $A \subset \mathbb{R}^n$ be closed and convex. A point $x \in A$ is called *extreme point*, if x cannot be represented as a nontrivial convex combination of points of A, i.e. if $x = \alpha y + (1-\alpha)z$ with $y, z \in A$, $\alpha \in (0,1)$, implies that $x = y = z$. The set of all extreme points of A is denoted by $\operatorname{ext} A$.

Remarks. (1) If A is a closed half-space, $\operatorname{ext} A = \emptyset$. In general, $\operatorname{ext} A \neq \emptyset$, if and only if A does not contain any lines.

(2) For $x \in A$, we have $x \in \operatorname{ext} A$, if and only if $A \setminus \{x\}$ is convex. In fact, assume that $x \in \operatorname{ext} A$. Let $y, z \in A \setminus \{x\}$. Then $[y,z] \subset A$. If $[y,z] \not\subset A \setminus \{x\}$, then $x \in (y,z)$ which contradicts $x \in \operatorname{ext} A$. Hence $[y,z] \subset A \setminus \{x\}$, i.e. $A \setminus \{x\}$ is convex. Conversely, assume that $A \setminus \{x\}$ is convex. Let $y, z \in A$ and $\alpha \in (0,1)$ such that $x = \alpha y + (1-\alpha)z$. If $y \neq x$ and $z \neq x$, then $y, z \in A \setminus \{x\}$ and therefore $x \in [y,z] \subset A \setminus \{x\}$, a contradiction. Therefore, $y = x$ or $z = x$, which implies that $x = y = z$.

(3) For a polytope P, the preceding remark yields that $\operatorname{ext} P = \operatorname{vert} P$.

(4) If $\{x\}$ is a support set of A, then $x \in \operatorname{ext} A$. The converse is false, as the following example of a planar set A shows. A is the sum of a circle and a segment, each of the points x_i is extreme, but $\{x_i\}$ is not a support set.



The preceding remark explains why the following definition is relevant.

Definition. Let $A \subset \mathbb{R}^n$ be closed and convex. A point $x \in A$ is called *exposed point*, if $\{x\}$ is a support set of A. The set of all exposed points of A is denoted by $\exp A$.

Remark. In view of Remark (4) above, we have $\exp A \subset \operatorname{ext} A$.

Theorem 1.5.1 (MINKOWSKI). *Let $K \subset \mathbb{R}^n$ be compact and convex, and let $A \subset K$. Then, $K = \operatorname{conv} A$, if and only if $\operatorname{ext} K \subset A$. In particular, $K = \operatorname{conv} \operatorname{ext} K$.*

Proof. Suppose $K = \operatorname{conv} A$ and $x \in \operatorname{ext} K$. Assume $x \notin A$. Then $A \subset K \setminus \{x\}$. Since $K \setminus \{x\}$ is convex, $K = \operatorname{conv} A \subset K \setminus \{x\}$, a contradiction.

In the other direction, we need only show that $K = \operatorname{conv} \operatorname{ext} K$. We prove this by induction on n. For $n = 1$, a compact convex subset of $\mathbb{R}^1$ is a segment $[a, b]$ and $\operatorname{ext} [a, b] = \{a, b\}$.

Let $n \geq 2$ and suppose the result holds in dimension $n - 1$. Since $\operatorname{ext} K \subset K$, we obviously have $\operatorname{conv} \operatorname{ext} K \subset K$. We need to show the opposite inclusion. For that purpose, let $x \in K$ and g an arbitrary line through x. Then $g \cap K = [y, z]$ with $x \in [y, z]$ and $y, z \in \operatorname{bd} K$. By the support theorem, y, z are support points, i.e. there are supporting hyperplanes E_y, E_z of K with $y \in K_1 := E_y \cap K$ and $z \in K_2 := E_z \cap K$. By the induction hypothesis,

$$K_1 = \operatorname{conv} \operatorname{ext} K_1; \quad K_2 = \operatorname{conv} \operatorname{ext} K_2.$$

We have $\operatorname{ext} K_1 \subset \operatorname{ext} K$. Namely, consider $u \in \operatorname{ext} K_1$ and $u = \alpha v + (1 - \alpha) w$, $v, w \in K$, $\alpha \in (0, 1)$. Since u lies in the supporting hyperplane E_y, the same must hold for v and w. Hence $v, w \in K_1$ and since $u \in \operatorname{ext} K_1$, we obtain $u = v = w$. Therefore, $u \in \operatorname{ext} K$.

In the same way, we get $\operatorname{ext} K_2 \subset \operatorname{ext} K$ and thus

$$\begin{aligned} x \in [y, z] &\subset \operatorname{conv} \{\operatorname{conv} \operatorname{ext} K_1 \cup \operatorname{conv} \operatorname{ext} K_2\} \\ &\subset \operatorname{conv} \operatorname{ext} K. \end{aligned}$$

□

Corollary 1.5.2. *Let $P \subset \mathbb{R}^n$ be compact and convex. Then P is a polytope, if and only if* $\operatorname{ext} P$ *is finite.*

Proof. If P is a polytope, then Theorem 1.1.5 and the preceding Remark (3) show that $\operatorname{ext} P$ is finite. For the converse, assume that $\operatorname{ext} P$ is finite, hence $\operatorname{ext} P = \{x_1, \ldots, x_k\}$. Theorem 1.5.1 then shows $P = \operatorname{conv} \{x_1, \ldots, x_k\}$, hence P is a polytope. □

Now we are able to prove a converse of Corollary 1.4.9.

Theorem 1.5.3. *Let $P \subset \mathbb{R}^n$ be a bounded polyhedral set. Then P is a polytope.*

Proof. Clearly, P is compact and convex. We show that $\operatorname{ext} P$ is finite.

Let $x \in \operatorname{ext} P$ and assume $P = \bigcap_{i=1}^k H_i$ with half-spaces H_i bounded by the hyperplanes E_i, $i = 1, \ldots, k$. We consider the convex set

$$D := \bigcap_{i=1}^{k} A_i,$$

where

$$A_i = \begin{cases} E_i & \text{if} \quad x \in E_i\,, \\ \operatorname{int} H_i & \quad x \notin E_i\,. \end{cases}$$

Then $x \in D \subset P$. Since x is an extreme point and D is relatively open, we get $\dim D = 0$, hence $D = \{x\}$. Since there are only finitely many different sets D possible, $\operatorname{ext} P$ must be finite. The result now follows from Corollary 1.5.2. □

Remark. This result now shows that the intersection of finitely many polytopes is again a polytope.

If we replace, in Theorem 1.5.1, the set $\operatorname{ext} K$ by $\exp K$, the corresponding result will be wrong, in general, as simple examples show (compare Theorem 1.5.1 and the remark preceding it). There is however a modified version which holds true for exposed points.

Theorem 1.5.4. *Let $K \subset \mathbb{R}^n$ be compact and convex. Then*

$$K = \operatorname{cl}\operatorname{conv}\,\exp K.$$

Proof. Since K is compact, for each $x \in \mathbb{R}^n$ there exists a point $y_x \in K$ farthest away from x, i.e. a point with

$$\|y_x - x\| = \max_{y \in K} \|y - x\|.$$

The hyperplane E through y_x orthogonal to $y_x - x$ is then a supporting hyperplane of K and we have $E \cap K = \{y_x\}$, hence $y_x \in \exp K$. Let

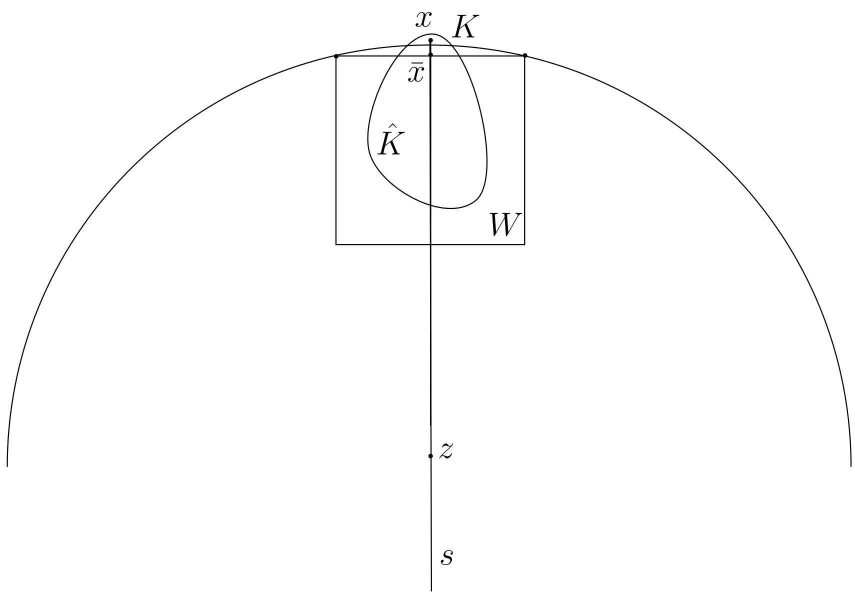
$$\hat{K} := \operatorname{cl}\operatorname{conv}\{y_x : x \in \mathbb{R}^n\}.$$

Then $\hat{K} \subset K$, thus $\hat{K}$ is compact.

Assume that there exists $x \in K \setminus \hat{K}$. Then, by Theorem 1.4.2 there is a hyperplane $E' = \{f = \alpha\}$ with $x \in \{f > \alpha\}$ and $\hat{K} \subset \{f \leq \alpha\}$ (E' is the supporting hyperplane through $p(\hat{K}, x)$ in direction $x - p(\hat{K}, x)$). Consider the half-line s starting in x, orthogonal to E' and in direction of that half-space of E', which contains $\hat{K}$. On s, we can find a point z with

$$\|x - z\| > \max_{y \in \hat{K}} \|y - z\|.$$

In fact, we may choose a cube W large enough to contain $\hat{K}$, and such that $p(\hat{K}, x)$ is the center of a facet of W. Now we choose a ball B with center $z \in s$ in such a way that $W \subset B$, but $x \notin B$. Then z is the required point.



By definition of $\hat{K}$, there exists $y_z \in \hat{K}$ with

$$\|y_z - z\| = \max_{y \in K} \|y - z\| \geq \|x - z\|,$$

a contradiction. Therefore, $K = \hat{K}$. Because of $y_x \in \exp K$, for all $x \in \mathbb{R}^n$, we obtain

$$K = \hat{K} \subset \operatorname{cl} \operatorname{conv} \exp K \subset K,$$

hence $K = \operatorname{cl} \operatorname{conv} \exp K$. □

Corollary 1.5.5 (STRASZEWICZ)**.** *Let $K \subset \mathbb{R}^n$ be compact and convex. Then*

$$\operatorname{ext} K \subset \operatorname{cl}\ \exp K.$$

Proof. By Theorems 1.5.4 and 1.3.7, we have

$$K = \operatorname{cl} \operatorname{conv} \exp K \subset \operatorname{cl} \operatorname{conv} \operatorname{cl} \exp K = \operatorname{conv} \operatorname{cl} \exp K \subset K\ ,$$

hence

$$K = \operatorname{conv} \operatorname{cl} \exp K.$$

By Theorem 1.5.1, this implies $\operatorname{ext} K \subset \operatorname{cl}\ \exp K$. □

Exercises and problems

1. Let $A \subset \mathbb{R}^n$ be closed and convex. Show that $\mathrm{ext}\, A \neq \emptyset$, if and only if A does not contain any line.

2. Let $K \subset \mathbb{R}^n$ be compact and convex.

   (a) If $n = 2$, show that $\mathrm{ext}\, K$ is closed.

   (b) If $n \geq 3$, show by an example that $\mathrm{ext}\, K$ need not be closed.

3. Let $A \subset \mathbb{R}^n$ be closed and convex. A subset $M \subset A$ is called *extreme set* (in A), if M is convex and if $x, y \in A$, $(x, y) \cap M \neq \emptyset$ implies $[x, y] \subset M$.

   Show that:

   (a) Extreme sets M are closed.

   (b) Each support set of A is extreme.

   (c) If $M, N \subset A$ are extreme, then $M \cap N$ is extreme.

   (d) If M is extreme in A and $N \subset M$ is extreme in M, then N is extreme in A.

   (e) If $M, N \subset A$ are extreme and $M \neq N$, then $\mathrm{relint}\, M \cap \mathrm{relint}\, N = \emptyset$.

   (f) Let $\mathcal{E}(A) := \{M \subset A :\ M \text{ extreme}\}$. Then $A = \bigcup_{M \in \mathcal{E}(A)} \mathrm{relint}\, M$ is a disjoint union.

4. A real (n, n)-matrix $A = ((\alpha_{ij}))$ is called *doubly stochastic*, if $\alpha_{ij} \geq 0$ and

$$\sum_{k=1}^{n} \alpha_{kj} = \sum_{k=1}^{n} \alpha_{ik} = 1$$

   for all $i, j \in \{1, \dots, n\}$. A doubly stochastic matrix with components in $\{0, 1\}$ is called *permutation matrix.*

   Show:

   (a) The set $K \subset \mathbb{R}^{n^2}$ of doubly stochastic matrices is compact and convex.

   (b) The extreme points of K are precisely the permutation matrices.
   Hint for (b): You may use the following simple combinatorial result (marriage theorem): Given a finite set H, a nonempty set D and a function $f : H \to \mathcal{P}(D)$ with

$$\Big| \bigcup_{h \in \tilde{H}} f(h) \Big| \geq |\tilde{H}|, \quad \text{for all } \tilde{H} \subset H,$$

   then there exists an injective function $g : H \to D$ with $g(h) \in f(h)$, for all $h \in H$.

Chapter 2

Convex functions

2.1 Properties and operations of convex functions

In the following, we consider functions

$$f : \mathbb{R}^n \to [-\infty, \infty].$$

We assume the usual rules for addition and multiplication with ∞, namely:

$$\begin{array}{ll} \alpha + \infty := \infty, & \text{for } \alpha \in (-\infty, \infty], \\ \alpha - \infty := -\infty, & \text{for } \alpha \in [-\infty, \infty), \\ \alpha\,\infty := \infty,\ (-\alpha)\infty := -\infty, & \text{for } \alpha \in (0, \infty], \\ 0\,\infty := 0. & \end{array}$$

Definition. For a function $f : \mathbb{R}^n \to (-\infty, \infty]$, the set

$$\operatorname{epi} f := \{(x, \alpha) : x \in \mathbb{R}^n, \alpha \in \mathbb{R}, f(x) \leq \alpha\} \subset \mathbb{R}^n \times \mathbb{R}$$

is called the *epigraph* of f. f is *convex*, if $\operatorname{epi} f$ is a convex subset of $\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$.

Remarks. (1) A function $f : \mathbb{R}^n \to [-\infty, \infty)$ is *concave*, if $-f$ is convex. Thus, for a convex function f we exclude the value $-\infty$, whereas for a concave function we exclude ∞.

(2) If $A \subset \mathbb{R}^n$ is a subset, a function $f : A \to (-\infty, \infty)$ is called *convex*, if the extended function $\tilde{f} : \mathbb{R}^n \to (-\infty, \infty]$, given by

$$\tilde{f} := \begin{cases} f & \text{on} \quad A\,, \\ \infty & \quad \mathbb{R}^n \setminus A\,, \end{cases}$$

is convex. This automatically requires that A is a convex set. In view of this construction, we need not consider convex functions defined on subsets of $\mathbb{R}^n$, but we rather can assume that convex functions are always defined on all of $\mathbb{R}^n$.

(3) On the other hand, we often are only interested in convex functions $f : \mathbb{R}^n \to (-\infty, \infty]$ at points, where f is finite. We call

$$\operatorname{dom} f := \{x \in \mathbb{R}^n : f(x) < \infty\}$$

the *effective domain* of the function $f : \mathbb{R}^n \to (-\infty, \infty]$. For a convex function f, the effective domain $\operatorname{dom} f$ is convex.

(4) The function $f \equiv \infty$ is convex, it is called the *improper* convex function; convex functions f with $f \not\equiv \infty$ are called *proper*. The improper convex function $f \equiv \infty$ has $\operatorname{epi} f = \emptyset$ and $\operatorname{dom} f = \emptyset$.

Theorem 2.1.1. *A function $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex, if and only if*

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y),$$

for all $x, y \in \mathbb{R}^n, \alpha \in [0,1]$.

Proof. By definition, f is convex, if and only if $\operatorname{epi} f = \{(x,\beta) : f(x) \le \beta\}$ is convex. The latter condition means

$$\alpha(x_1,\beta_1) + (1-\alpha)(x_2,\beta_2) = (\alpha x_1 + (1-\alpha)x_2, \alpha\beta_1 + (1-\alpha)\beta_2) \in \operatorname{epi} f,$$

for all $\alpha \in [0,1]$ and whenever $(x_1,\beta_1), (x_2,\beta_2) \in \operatorname{epi} f$, i.e. whenever $f(x_1) \le \beta_1$, $f(x_2) \le \beta_2$.

Hence, f is convex, if and only if

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha\beta_1 + (1-\alpha)\beta_2,$$

for all $x_1, x_2 \in \mathbb{R}^n, \alpha \in [0,1]$ and all $\beta_1 \ge f(x_1), \beta_2 \ge f(x_2)$. Then, it is necessary and sufficient that this inequality is satisfied for $\beta_1 = f(x_1), \beta_2 = f(x_2)$, and we obtain the assertion. □

Remarks. (1) A function $f : \mathbb{R}^n \to \mathbb{R}$ is affine, if and only if f is convex and concave. If f is affine, then $\operatorname{epi} f$ is a half-space in $\mathbb{R}^{n+1}$ (and $\operatorname{dom} f = \mathbb{R}^n$).

(2) For a convex function f, the sublevel sets $\{f < \alpha\}$ and $\{f \le \alpha\}$ are convex.

(3) If f, g are convex and $\alpha, \beta \ge 0$, then $\alpha f + \beta g$ is convex.

(4) If $(f_i)_{i\in I}$ is a family of convex functions, the (pointwise) supremum $\sup_{i\in I} f_i$ is convex. This follows since

$$\operatorname{epi}\left(\sup_{i\in I} f_i\right) = \bigcap_{i\in I} \operatorname{epi} f_i.$$

(5) As a generalization of Theorem 2.1.1, we obtain that f is convex, if and only if

$$f(\alpha_1 x_1 + \cdots + \alpha_k x_k) \le \alpha_1 f(x_1) + \cdots + \alpha_k f(x_k),$$

for all $k \in \mathbb{N}$, $x_i \in \mathbb{R}^n$, and $\alpha_i \in [0,1]$ with $\sum \alpha_i = 1$.

(6) A function $f : \mathbb{R}^n \to (-\infty, \infty]$ is *positively homogeneous* (of degree 1), if

$$f(\alpha x) = \alpha f(x), \quad \text{for all } x \in \mathbb{R}^n, \alpha \ge 0.$$

If f is positively homogeneous, f is convex if and only if it is *subadditive*, i.e. if

$$f(x+y) \le f(x) + f(y),$$

for all $x, y \in \mathbb{R}^n$.

The following simple result is useful for generating convex functions from convex sets in $\mathbb{R}^n \times \mathbb{R}$.

Theorem 2.1.2. *Let $A \subset \mathbb{R}^n \times \mathbb{R}$ be convex and suppose that*

$$f_A(x) := \inf\{\alpha \in \mathbb{R} : (x,\alpha) \in A\} > -\infty,$$

for all $x \in \mathbb{R}^n$. Then, f_A is a convex function.

Proof. The definition of $f_A(x)$ implies that

$$\operatorname{epi} f_A = \{(x,\beta) : \exists \alpha \in \mathbb{R}, \alpha \leq \beta, \text{ and a sequence } \alpha_i \searrow \alpha \text{ with } (x,\alpha_i) \in A\}.$$

It is easy to see that $\operatorname{epi} f_A$ is convex. □

Remarks. (1) The condition $f_A > -\infty$ is fulfilled, if and only if A does not contain a vertical half-line which is unbounded from below.
(2) For $x \in \mathbb{R}^n$, let

$$\{x\} \times \mathbb{R} := \{(x,\alpha) : \alpha \in \mathbb{R}\}$$

be the *vertical line* in $\mathbb{R}^n \times \mathbb{R}$ through x. Let $A \subset \mathbb{R}^n \times \mathbb{R}$ be closed and convex. Then, we have $A = \operatorname{epi} f_A$, if and only if

$$A \cap (\{x\} \times \mathbb{R}) = \{x\} \times [f_A(x), \infty), \quad \text{for all } x \in \mathbb{R}^n.$$

Theorem 2.1.2 allows us to define operations of convex functions by applying corresponding operations of convex sets to the epigraphs of the functions. We give two examples of that kind.

Definition. A convex function $f : \mathbb{R}^n \to (-\infty, \infty]$ is *closed*, if $\operatorname{epi} f$ is closed.

If $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex, then $\operatorname{cl} \operatorname{epi} f$ is the epigraph of a closed convex function, which we denote by $\operatorname{cl} f$.

To see this, we have to show that $A := \operatorname{cl} \operatorname{epi} f$ fulfills $f_A > -\infty$. The case $f \equiv \infty$ is trivial, then f is closed and $f_A = f$.

Let f be proper, then $\operatorname{epi} f \neq \emptyset$. W.l.o.g. we may assume that $\dim \operatorname{dom} f = n$. We choose a point $x \in \operatorname{int} \operatorname{dom} f$. Then, $(x, f(x)) \in \operatorname{bd} \operatorname{epi} f$. Hence, there is a supporting hyperplane $E \subset \mathbb{R}^n \times \mathbb{R}$ of $\operatorname{cl} \operatorname{epi} f$ at $(x, f(x))$. The corresponding supporting half-space is the epigraph of an affine function $h \leq f$. Thus, $f_A \geq h > -\infty$.

Remark. $\operatorname{cl} f$ is the largest closed convex function below f.

Our second example is the convex hull operator. If $(f_i)_{i \in I}$ is a family of (arbitrary) functions $f_i : \mathbb{R}^n \to (-\infty, \infty]$, we consider $A := \bigcup_{i \in I} \operatorname{epi} f_i$. Suppose $\operatorname{conv} A$ does not contain any vertical line, then, by Theorem 2.1.2, $\operatorname{conv}(f_i) := f_{\operatorname{conv} A}$ is a convex function, which we call the *convex hull* of the functions $f_i, i \in I$. It is easy to see, that $\operatorname{conv}(f_i)$ is the largest convex function below all f_i, i.e.

$$\operatorname{conv}(f_i) = \sup\{g : g \text{ convex}, \ g \leq f_i \ \forall i \in I\}.$$

$\operatorname{conv}(f_i)$ exists, if and only if there is an affine function h with $h \leq f_i$, for all $i \in I$.

Further applications of Theorem 2.1.2 are listed in the exercises.

The following representation of convex functions is a counterpart to the support theorem for convex sets.

Theorem 2.1.3. *Let* $f : \mathbb{R}^n \to (-\infty, \infty]$ *be closed and convex. Then,*

$$f = \sup \{h : h \leq f, h \text{ affine}\}.$$

Proof. By assumption, $\operatorname{epi} f$ is closed and convex. Moreover, we can assume that f is proper, i.e. $\operatorname{epi} f \neq \emptyset$. By Corollary 1.4.3, $\operatorname{epi} f$ is the intersection of all closed half-spaces $H \subset \mathbb{R}^n \times \mathbb{R}$ which contain $\operatorname{epi} f$.

There are three types of closed half-spaces in $\mathbb{R}^n \times \mathbb{R}$:

$$\begin{aligned} H_1 &= \{(x, r) : r \geq l(x)\}, \quad && l : \mathbb{R}^n \to \mathbb{R} \text{ affine}, \\ H_2 &= \{(x, r) : r \leq l(x)\}, \quad && l : \mathbb{R}^n \to \mathbb{R} \text{ affine}, \\ H_3 &= \tilde{H} \times \mathbb{R}, && \tilde{H} \text{ half-space in } \mathbb{R}^n. \end{aligned}$$

Half-spaces of type H_2 cannot occur, due to the definition of $\operatorname{epi} f$ and since $\operatorname{epi} f \neq \emptyset$. Half-spaces of type H_3 can occur, hence we have to show that these 'vertical' half-spaces can be avoided, i.e. $\operatorname{epi} f$ is the intersection of all half-spaces of type H_1 containing $\operatorname{epi} f$. Then we are finished since the intersection of half-spaces of type H_1 is the epigraph of the supremum of the corresponding affine functions l.

For the result just explained it is sufficient to show that any point $(x_0, r_0) \notin \operatorname{epi} f$ can be separated by a non-vertical hyperplane E from $\operatorname{epi} f$. Hence, let E_3 be a vertical hyperplane separating (x_0, r_0) and $\operatorname{epi} f$, obtained from Theorem 1.4.2, and let H_3 be the corresponding vertical half-space containing $\operatorname{epi} f$. Since $f > -\infty$, there is at least one affine function l_1 with $l_1 \leq f$. We may represent H_3 as

$$H_3 = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : l_0(x) \leq 0\}$$

with some affine function $l_0 : \mathbb{R}^n \to \mathbb{R}$, and we may assume $l_0(x_0) > 0$.

For $x \in \operatorname{dom} f$, we then have

$$l_0(x) \leq 0, \quad l_1(x) \leq f(x),$$

hence

$$\alpha l_0(x) + l_1(x) \leq f(x), \quad \text{for all } \alpha \geq 0.$$

For $x \notin \operatorname{dom} f$, this inequality holds trivially since then $f(x) = \infty$. Hence

$$m_\alpha := \alpha l_0 + l_1$$

is an affine function fulfilling $m_\alpha \leq f$. Since $l_0(x_0) > 0$, we have $m_\alpha(x_0) > r_0$ for sufficiently large α. □

We now come to another important operation on convex functions, the construction of the conjugate function.

Definition. Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be proper and convex, then the function f^* defined by

$$f^*(y) := \sup_{x \in \mathbb{R}^n} (\langle x, y \rangle - f(x)), \quad y \in \mathbb{R}^n,$$

is called the *conjugate* of f.

Theorem 2.1.4. *The conjugate f^* of a proper convex function $f : \mathbb{R}^n \to (-\infty, \infty]$ fulfills:*
(a) *f^* is proper, closed and convex.*
(b) $f^{**} := (f^*)^* = \operatorname{cl} f$.

Proof. (a) For $x \notin \operatorname{dom} f$, we have $\langle x, y\rangle - f(x) = -\infty$ (for all $y \in \mathbb{R}^n$), hence

$$f^* = \sup_{x \in \operatorname{dom} f} (\langle x, \cdot\rangle - f(x)).$$

For $x \in \operatorname{dom} f$, the function

$$g_x : y \mapsto \langle x, y\rangle - f(x)$$

is affine, therefore f^* is convex (as the supremum of affine functions).

Because of

$$\operatorname{epi} f^* = \operatorname{epi} \left(\sup_{x \in \operatorname{dom} f} g_x \right) = \bigcap_{x \in \operatorname{dom} f} \operatorname{epi} g_x$$

and since $\operatorname{epi} g_x$ is a closed half-space, $\operatorname{epi} f^*$ is closed, and hence f^* is closed.

In order to show that f^* is proper, we consider an affine function $h \leq f$. Such a function exists by Theorem 2.1.3 and it has a representation

$$h = \langle \cdot, y\rangle - \alpha, \quad \text{with suitable } y \in \mathbb{R}^n, \alpha \in \mathbb{R}.$$

This implies

$$\langle \cdot, y\rangle - \alpha \leq f,$$

hence

$$\langle \cdot, y\rangle - f \leq \alpha,$$

and therefore $f^*(y) \leq \alpha$.

(b) By Theorem 2.1.3,

$$\operatorname{cl} f = \sup\{h : h \leq \operatorname{cl} f, h \text{ affine}\}.$$

Writing h again as

$$h = \langle \cdot, y\rangle - \alpha, \quad y \in \mathbb{R}^n, \alpha \in \mathbb{R},$$

we obtain

$$\operatorname{cl} f = \sup_{(y,\alpha)} (\langle \cdot, y\rangle - \alpha),$$

where the supremum is taken over all (y, α) with

$$\langle \cdot, y\rangle - \alpha \leq \operatorname{cl} f.$$

The latter holds, if and only if

$$\alpha \geq \sup_x (\langle x, y\rangle - \operatorname{cl} f(x)) = (\operatorname{cl} f)^*(y).$$

Consequently, we have

$$\operatorname{cl} f(x) \le \sup_y(\langle x, y\rangle - (\operatorname{cl} f)^*(y)) = (\operatorname{cl} f)^{**}(x),$$

for all x. Since $\operatorname{cl} f \le f$, the definition of the conjugate function implies

$$(\operatorname{cl} f)^* \ge f^*,$$

and therefore

$$\operatorname{cl} f \le (\operatorname{cl} f)^{**} \le f^{**}.$$

On the other hand,

$$f^{**}(x) = (f^*)^*(x) = \sup_y(\langle x, y\rangle - f^*(y)),$$

where

$$f^*(y) = \sup_z(\langle z, y\rangle - f(z)) \ge \langle x, y\rangle - f(x).$$

Therefore,

$$f^{**}(x) \le \sup_y(\langle x, y\rangle - \langle x, y\rangle + f(x)) = f(x),$$

which gives us $f^{**} \le f$. By part (a), f^{**} is closed, hence $f^{**} \le \operatorname{cl} f$. □

Finally, we mention a canonical possibility to describe convex sets $A \subset \mathbb{R}^n$ by convex functions. The common way to describe a set A is by the function

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{if} \quad x \in A\,, \\ 0 & \phantom{\text{if}} \quad x \notin A\,, \end{cases}$$

however, $\mathbf{1}_A$ is neither convex nor concave. Therefore, we here define the *indicator function* δ_A of a (arbitrary) set $A \subset \mathbb{R}^n$ by

$$\delta_A(x) := \begin{cases} 0 & \text{if} \quad x \in A\,, \\ \infty & \phantom{\text{if}} \quad x \notin A\,. \end{cases}$$

Remark. A is convex, if and only if δ_A is convex.

Exercises and problems

1. Let $A \subset \mathbb{R}^n$ be nonempty, closed and convex and containing no line. Let further $f : \mathbb{R}^n \to \mathbb{R}$ be convex and assume there is a point $y \in A$ with

$$f(y) = \max_{x \in A} f(x).$$

   Show that there is also a $z \in \operatorname{ext} A$ with

$$f(z) = \max_{x \in A} f(x).$$

2. Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be convex. Show that the following assertions are equivalent:
   (i) f is closed.
   (ii) f is lower semi-continuous, i.e. for all $x \in \mathbb{R}^n$ we have
   $$f(x) \leq \liminf_{y \to x} f(y).$$
   (iii) All the sublevel sets $\{f \leq \alpha\}$, $\alpha \in \mathbb{R}$, are closed .
3. Let $f, f_1, \dots, f_m : \mathbb{R}^n \to (-\infty, \infty]$ be convex functions and $\alpha \geq 0$. Show that:
   (a) The function $\alpha \circ f : x \mapsto \inf\{\beta \in \mathbb{R} : (x, \beta) \in \alpha \cdot \operatorname{epi} f\}$ is convex.
   (b) The function $f_1 \,\square \cdots \square\, f_m : x \mapsto \inf\{\beta \in \mathbb{R} : (x, \beta) \in \operatorname{epi} f_1 + \cdots + \operatorname{epi} f_m\}$ is convex, and we have
   $$f_1 \,\square \cdots \square\, f_m(x) = \inf\{f_1(x_1) + \cdots + f_m(x_m) : x_1, \dots, x_m \in \mathbb{R}^n, x_1 + \cdots + x_m = x\}.$$
   ($f_1 \,\square \cdots \square\, f_m$ is called the *infimal convolution* of $f_1, \dots, f_m$.)
   (c) Let $\{f_i : i \in I\}$ $(I \neq \emptyset)$ be a family of convex functions on $\mathbb{R}^n$, such that $\operatorname{conv}(f_i)$ exists. Show that
   $$\operatorname{conv}(f_i) = \inf\{\alpha_1 \circ f_{i_1} \,\square \cdots \square\, \alpha_m \circ f_{i_m} : \alpha_j \geq 0, \sum \alpha_j = 1, i_j \in I, m \in \mathbb{N}\}.$$
4. Let $A \subset \mathbb{R}^n$ be convex and $0 \in A$. The *distance function* $d_A : \mathbb{R}^n \to (-\infty, \infty]$ is defined as
   $$d_A(x) = \inf\{\alpha \geq 0 : x \in \alpha A\}, \quad x \in \mathbb{R}^n.$$
   Show that d_A has the following properties:
   (a) d_A is positively homogeneous, nonnegative and convex.
   (b) d_A is finite, if and only if $0 \in \operatorname{int} A$.
   (c) $\{d_A < 1\} \subset A \subset \{d_A \leq 1\} \subset \operatorname{cl} A$.
   (d) If $0 \in \operatorname{int} A$, then $\operatorname{int} A = \{d_A < 1\}$ and $\operatorname{cl} A = \{d_A \leq 1\}$.
   (e) $d_A(x) > 0$, if and only if $x \neq 0$ and $\beta x \notin A$ for some $\beta > 0$.
   (f) Let A be closed. Then d_A is even (i.e. $d_A(x) = d_A(-x) \, \forall x \in \mathbb{R}^n$), if and only if A is symmetric with respect to 0 (i.e. $A = -A$).
   (g) Let A be closed. Then d_A is a norm on $\mathbb{R}^n$, if and only if A is symmetric, compact and contains 0 in its interior.
   (h) If A is closed, then d_A is closed.

2.2 Regularity of convex functions

We start with a continuity property of convex functions.

Theorem 2.2.1. *A convex function* $f : \mathbb{R}^n \to (-\infty, \infty]$ *is continuous in* $\operatorname{int} \operatorname{dom} f$ *and Lipschitz continuous on compact subsets of* $\operatorname{int} \operatorname{dom} f$.

Proof. Let $x \in \operatorname{int} \operatorname{dom} f$. There exists a n-simplex P with $P \subset \operatorname{int} \operatorname{dom} f$ and $x \in \operatorname{int} P$. If $x_0, \dots, x_n$ are the vertices of P and $y \in P$, we have

$$y = \alpha_0 x_0 + \cdots + \alpha_n x_n,$$

with $\alpha_i \in [0,1]$, $\sum \alpha_i = 1$, and hence

$$f(y) \le \alpha_0 f(x_0) + \cdots + \alpha_n f(x_n) \le \max_{i=0,\dots,n} f(x_i) =: c.$$

Therefore, $f \le c$ on P.

Let now $\alpha \in (0,1)$ and choose an open ball U centered at 0 such that $x + U \subset P$. Let $z = x + \alpha u, u \in \operatorname{bd} U$. Then,

$$z = (1-\alpha)x + \alpha(x+u),$$

$$f(z) \le (1-\alpha) f(x) + \alpha f(x+u) \le (1-\alpha) f(x) + \alpha C,$$

where $C := \max\{|f(y)| : y \in x + \operatorname{cl} U\} \le c$. This gives us

$$f(z) - f(x) \le \alpha(C - f(x)).$$

On the other hand,

$$x = \frac{1}{1+\alpha}(x + \alpha u) + \left(1 - \frac{1}{1+\alpha}\right)(x-u),$$

and hence

$$f(x) \le \frac{1}{1+\alpha} f(x + \alpha u) + \left(1 - \frac{1}{1+\alpha}\right) f(x-u),$$

which implies

$$f(x) \le \frac{1}{1+\alpha} f(z) + \frac{\alpha}{1+\alpha} C.$$

We obtain

$$\alpha(f(x) - C) \le f(z) - f(x).$$

Together, the two inequalities give

$$|f(z) - f(x)| \le \alpha(C - f(x)),$$

for all $z \in x + \alpha U$. Let ϱ be the radius of U. Thus we have shown that

$$|f(z) - f(x)| \le \frac{2C}{\varrho} \|z - x\|.$$

Now let $A \subset \operatorname{int} \operatorname{dom} f$ be compact. Hence there is some $\varrho > 0$ such that $A + \varrho B^n \subset \operatorname{int} \operatorname{dom} f$. Let $x, z \in A$. Since f is continuous on $A + \varrho B^n$,

$$\tilde{C} := \max\{|f(y)| : y \in A + \varrho B^n\} < \infty.$$

By the preceding argument,

$$|f(z) - f(x)| \leq \frac{2\tilde{C}}{\varrho}\|z - x\|,$$

if $\|z - x\| \leq \varrho$. For $\|z - x\| \geq \varrho$, this is true as well. □

Now we discuss differentiability properties of convex functions. We first consider the case $f : \mathbb{R}^1 \to (-\infty, \infty]$.

Theorem 2.2.2. *Let $f : \mathbb{R}^1 \to (-\infty, \infty]$ be convex.*

(a) *In each point $x \in \operatorname{int} \operatorname{dom} f$, the right derivative $f^+(x)$ and the left derivative $f^-(x)$ exist and fulfill $f^-(x) \leq f^+(x)$.*

(b) *On $\operatorname{int} \operatorname{dom} f$, the functions f^+ and f^- are monotonically increasing and, for almost all $x \in \operatorname{int} \operatorname{dom} f$ (with respect to the Lebesgue measure λ_1 on $\mathbb{R}^1$), we have $f^-(x) = f^+(x)$, hence f is almost everywhere differentiable on $\operatorname{cl} \operatorname{dom} f$.*

(c) *Moreover, f^+ is continuous from the right and f^- continuous from the left, and f is the indefinite integral of f^+ (of f^- and of f') in $\operatorname{int} \operatorname{dom} f$.*

Proof. W.l.o.g. we concentrate on the case $\operatorname{dom} f = \mathbb{R}^1$.

(a) If $0 < m \leq l$ and $0 < h \leq k$, the convexity of f implies

$$f(x - m) = f((1 - \frac{m}{l})x + \frac{m}{l}(x - l)) \leq (1 - \frac{m}{l})f(x) + \frac{m}{l}f(x - l),$$

hence

$$\frac{f(x) - f(x - l)}{l} \leq \frac{f(x) - f(x - m)}{m}.$$

Similarly, we have

$$f(x) = f(\frac{h}{h + m}(x - m) + \frac{m}{h + m}(x + h)) \leq \frac{h}{h + m}f(x - m) + \frac{m}{h + m}f(x + h),$$

which gives us

$$\frac{f(x) - f(x - m)}{m} \leq \frac{f(x + h) - f(x)}{h}.$$

Finally,

$$f(x + h) = f((1 - \frac{h}{k})x + \frac{h}{k}(x + k)) \leq (1 - \frac{h}{k})f(x) + \frac{h}{k}f(x + k),$$

and therefore

$$\frac{f(x + h) - f(x)}{h} \leq \frac{f(x + k) - f(x)}{k}.$$

We obtain that the left difference quotients in x increase monotonically and are bounded above by the right difference quotients, which decrease monotonically. Therefore, the limits

$$f^+(x) = \lim_{h \searrow 0} \frac{f(x+h) - f(x)}{h}$$

and

$$f^-(x) = \lim_{m \searrow 0} \frac{f(x) - f(x-m)}{m} \qquad \left(= \lim_{t \nearrow 0} \frac{f(x+t) - f(x)}{t} \right)$$

exist and fulfill $f^-(x) \le f^+(x)$.

(b) For $x' > x$, we have just seen that

$$f^-(x) \le f^+(x) \le \frac{f(x') - f(x)}{x' - x} \le f^-(x') \le f^+(x'). \tag{2.1}$$

Therefore, the functions f^- and f^+ are monotonically increasing. As is well-known, a monotonically increasing function has only countably many points of discontinuity (namely jumps), and therefore it is continuous almost everywhere. In the points x of continuity of f^-, (2.1) implies $f^-(x) = f^+(x)$.

(c) Assume now $x < y$. From

$$\frac{f(y) - f(x)}{y - x} = \lim_{z \searrow x} \frac{f(y) - f(z)}{y - z} \ge \lim_{z \searrow x} f^+(z)$$

we obtain $\lim_{z \searrow x} f^+(z) \le f^+(x)$, hence $\lim_{z \searrow x} f^+(z) = f^+(x)$, since f^+ is increasing. For $y < x$, we get by a similar argument

$$\lim_{z \nearrow x} f^-(z) \ge \lim_{z \nearrow x} \frac{f(z) - f(y)}{z - y} = \frac{f(x) - f(y)}{x - y},$$

and hence $f^-(x) \le \lim_{z \nearrow x} f^-(z) \le f^-(x)$. Thus we also have $\lim_{z \nearrow x} f^-(z) = f^-(x)$.

Finally, for arbitrary $a \in \mathbb{R}$, we define a function g by

$$g(x) := f(a) + \int_a^x f^-(s)\, ds.$$

We first show that g is convex, and then $g = f$.

For $z := \alpha x + (1 - \alpha)y, \alpha \in [0,1], x < y$, we have

$$g(z) - g(x) = \int_x^z f^-(s)\, ds \le (z - x) f^-(z),$$

$$g(y) - g(z) = \int_z^y f^-(s)\, ds \ge (y - z) f^-(z).$$

It follows that

$$\begin{aligned} \alpha(g(z) - g(x)) + (1 - \alpha)(g(z) - g(y)) &\le \alpha(z - x) f^-(z) + (1 - \alpha)(z - y) f^-(z) \\ &= f^-(z)(z - [\alpha x + (1 - \alpha) y]) = 0, \end{aligned}$$

therefore

$$g(z) \le \alpha g(x) + (1-\alpha) g(y),$$

i.e. g is convex.

As a consequence, g^+ and g^- exist. For $y > x$,

$$\frac{g(y)-g(x)}{y-x} = \frac{1}{y-x}\int_x^y f^-(s)\,ds = \frac{1}{y-x}\int_x^y f^+(s)\,ds \ge f^+(x),$$

hence we obtain $g^+(x) \ge f^+(x)$. Analogously,

$$\frac{g(x)-g(y)}{x-y} = \frac{1}{x-y}\int_y^x f^-(s)\,ds \le f^-(x),$$

and thus we get $g^-(x) \le f^-(x)$. Since $g^+ \ge f^+ \ge f^- \ge g^-$ and $g^+ = g^-$, except for at most countably many points, we have $g^+ = f^+$ and $g^- = f^-$ except for at most countably many points. By the continuity from the left of g^- and f^-, and the continuity from the right of g^+ and f^+, it follows that $g^+ = f^+$ and $g^- = f^-$ on $\mathbb{R}$. Hence, $h := g - f$ is differentiable everywhere and $h' \equiv 0$. Therefore, $h \equiv c = 0$ because we have $g(a) = f(a)$. □

Now we consider the n-dimensional case. If $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex and $x \in \operatorname{int}\operatorname{dom} f$, then, for each $u \in \mathbb{R}^n, u \ne 0$, the equation

$$g_{(u)}(t) := f(x+tu), \qquad t \in \mathbb{R},$$

defines a convex function $g_{(u)} : \mathbb{R}^1 \to (-\infty, \infty]$ and we have $0 \in \operatorname{int}\operatorname{dom} g_{(u)}$. By Theorem 2.2.2, the right derivative $g^+_{(u)}(0)$ exists. This is precisely the *directional derivative*

$$f'(x;u) := \lim_{t \searrow 0} \frac{f(x+tu) - f(x)}{t} \tag{2.2}$$

of f in direction u. Therefore, we obtain the following corollary to Theorem 2.2.2.

Corollary 2.2.3. *Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be convex and $x \in \operatorname{int}\operatorname{dom} f$. Then, for each $u \in \mathbb{R}^n, u \ne 0$, the directional derivative $f'(x;u)$ of f exists.*

The corollary does not imply that $f'(x;u) = -f'(x;-u)$ holds (in fact, the latter equation is only true if $g^-_{(u)}(0) = g^+_{(u)}(0)$). Also, the partial derivatives $f_1(x), ..., f_n(x)$ of f need not exist in each point x. However, in analogy to Theorem 2.2.2, on can show that $f_1, ..., f_n$ exist almost everywhere (with respect to the Lebesgue measure λ_n in $\mathbb{R}^n$) and that in points x, where the partial derivatives $f_1(x), ..., f_n(x)$ exist, the function f is even differentiable. Even more, a convex function f on $\mathbb{R}^n$ is twice differentiable almost everywhere (in a suitable sense). We refer to the exercises, for these and a number of further results on derivatives of convex functions.

The right-hand side of (2.2) also makes sense for $u = 0$ and yields the value 0. We therefore define $f'(x;0) := 0$. Then $u \mapsto f'(x;u)$ is a positively homogeneous function on $\mathbb{R}^n$ and if f is convex, $f'(x;\cdot)$ is also convex. For support functions, we will continue the discussion of directional derivatives in the next section.

For a function f which is differentiable or twice differentiable, the first or second derivatives can be used to characterize convexity of f.

Remarks. (1) (see Exercise 3) Let $A \subset \mathbb{R}$ be open and convex and let $f : A \to \mathbb{R}$ be a real function.

If f is differentiable, then f is convex, if and only if f' is monotone increasing (on A).

If f is twice differentiable, then f is convex, if and only if $f'' \geq 0$ (on A).

(2) (see Exercise 4) Let $A \subset \mathbb{R}^n$ be open and convex and let $f : A \to \mathbb{R}$ be a real function.

If f is differentiable, then f is convex, if and only if

$$\langle \text{grad}\, f(x) - \text{grad}\, f(y), x - y \rangle \geq 0, \quad \text{for all } x, y \in A.$$

(Here, grad $f(x) := (f_1(x), \ldots, f_n(x))$ is the *gradient* of f at x.)

If f is twice differentiable, then f is convex, if and only if the *Hessian matrix*

$$\partial^2 f(x) := ((f_{ij}(x)))_{n \times n}$$

of f at x is positive semidefinite, for all $x \in A$.

Exercises and problems

1. (a) Give an example of two convex functions $f, g : \mathbb{R}^n \to (-\infty, \infty]$, such that f and g both have minimal points (i.e. points in $\mathbb{R}^n$, where the infimum of the function is attained), but $f + g$ does not have a minimal point.

   (b) Suppose $f, g : \mathbb{R}^n \to \mathbb{R}$ are convex functions, which both have a unique minimal point in $\mathbb{R}^n$. Show that $f + g$ has a minimal point.

   Hint: Show first that the sets

   $$\{x \in \mathbb{R}^n : f(x) \leq \alpha\} \qquad \text{resp.} \qquad \{x \in \mathbb{R}^n : g(x) \leq \alpha\}$$

   are compact, for each $\alpha \in \mathbb{R}$.

2. Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function. Show that

   $$f(x) - f(0) = \int_0^x f^+(t)\, dt = \int_0^x f^-(t)\, dt,$$

   for all $x \in \mathbb{R}$.

3. Let $A \subset \mathbb{R}$ be open and convex and $f : A \to \mathbb{R}$ a real function.

   (a) Assume f is differentiable. Show that f is convex, if and only if f' is monotone increasing (on A).

   (b) Assume f is twice differentiable. Show that f is convex, if and only if $f'' \geq 0$ (on A).

4. Let $A \subset \mathbb{R}^n$ be open and convex and $f : A \to \mathbb{R}$ a real function.

   (a) Assume f is differentiable. Show that f is convex, if and only if

   $$\langle \text{grad}\, f(x) - \text{grad}\, f(y), x - y \rangle \geq 0, \quad \text{for all } x, y \in A.$$

   (Here, $\text{grad}\, f(x) := (f_1(x), \ldots, f_n(x))$ is the *gradient* of f at x.)

   (b) Assume f is twice differentiable. Show that f is convex, if and only if the *Hessian matrix*

   $$\partial^2 f(x) := ((f_{ij}(x)))_{n \times n}$$

   of f at x is positive semidefinite, for all $x \in A$.

5. For a convex function $f : \mathbb{R}^n \to (-\infty, \infty]$ and $x \in \text{int}\,\text{dom}\, f$, we define the *subgradient* of f at x by

   $$\partial f(x) := \{v \in \mathbb{R}^n \,:\, f(y) \geq f(x) + \langle v, y - x \rangle \;\forall y \in \mathbb{R}^n\}.$$

   Show that:

   (a) $\partial f(x)$ is nonempty, compact and convex.

   (b) We have

   $$\partial f(x) = \{v \in \mathbb{R}^n \,:\, \langle v, u \rangle \leq f'(x; u) \;\; \forall u \in \mathbb{R}^n,\, u \neq 0\}.$$

   (c) If f is differentiable in x, then

   $$\partial f(x) = \{\text{grad}\, f(x)\}.$$

$*$ 6. Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be convex and $x \in \text{int}\,\text{dom}\, f$. Suppose that all partial derivatives $f_1(x), \ldots, f_n(x)$ at x exist. Show that f is differentiable at x.

7. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. Show that f is differentiable almost everywhere.

   Hint: Use Exercise 6.

2.3 The support function

The most useful analytic description of compact convex sets is by the support function. It is one of the basic tools in the following chapter. The support function of a set $A \subset \mathbb{R}^n$ with $0 \in A$ is in a certain sense dual to the distance function, which was discussed in Exercise 2.1.3.

Definition. Let $A \subset \mathbb{R}^n$ be nonempty and convex. The *support function* $h_A : \mathbb{R}^n \to (-\infty, \infty]$ of A is defined as

$$h_A(u) := \sup_{x \in A} \langle x, u \rangle, \quad u \in \mathbb{R}^n.$$

Theorem 2.3.1. *Let $A, B \subset \mathbb{R}^n$ be nonempty convex sets. Then*

(a) *h_A is positively homogeneous, closed and convex (and hence subadditive).*

(b) *$h_A = h_{\mathrm{cl}\, A}$ and*

$$\mathrm{cl}\, A = \{x \in \mathbb{R}^n : \langle x, u \rangle \leq h_A(u)\ \forall u \in \mathbb{R}^n\}.$$

(c) *$A \subset B$ implies $h_A \leq h_B$; conversely, $h_A \leq h_B$ implies $\mathrm{cl}\, A \subset \mathrm{cl}\, B$.*

(d) *h_A is finite, if and only if A is bounded.*

(e) *$h_{\alpha A + \beta B} = \alpha h_A + \beta h_B$, for all $\alpha, \beta \geq 0$.*

(f) *$h_{-A}(u) = h_A(-u)$, for all $u \in \mathbb{R}^n$.*

(g) *If $A_i, i \in I$, are nonempty and convex and $A := \mathrm{conv}\left(\bigcup_{i \in I} A_i\right)$, then*

$$h_A = \sup_{i \in I} h_{A_i}.$$

(h) *If $A_i, i \in I$, are nonempty, convex and closed and if $A := \bigcap_{i \in I} A_i$ is nonempty, then*

$$h_A = \mathrm{cl}\,\mathrm{conv}\,(h_{A_i})_{i \in I}.$$

(i) $\delta_A^* = h_A$.

Proof. (a) For $\alpha \geq 0$ and $u, v \in \mathbb{R}^n$, we have

$$h_A(\alpha u) = \sup_{x \in A} \langle x, \alpha u \rangle = \alpha \sup_{x \in A} \langle x, u \rangle = \alpha h_A(u)$$

and

$$h_A(u + v) = \sup_{x \in A} \langle x, u + v \rangle \leq \sup_{x \in A} \langle x, u \rangle + \sup_{x \in A} \langle x, v \rangle = h_A(u) + h_A(v).$$

Furthermore, as a supremum of closed functions, h_A is closed.

(b) The first part follows from

$$\sup_{x \in A} \langle x, u \rangle = \sup_{x \in \mathrm{cl}\, A} \langle x, u \rangle, \quad u \in \mathbb{R}^n.$$

For $x \in \operatorname{cl} A$, we therefore have $\langle x, u\rangle \leq h_A(u)$, for all $u \in \mathbb{R}^n$. Conversely, suppose $x \in \mathbb{R}^n$ fulfills $\langle x, \cdot\rangle \leq h_A(\cdot)$, and assume $x \notin \operatorname{cl} A$. Then, by Theorem 1.4.2, there exists a (supporting) hyperplane separating x and $\operatorname{cl} A$, i.e. a direction $y \in S^{n-1}$ and $\alpha \in \mathbb{R}$ such that

$$\langle x, y\rangle > \alpha \text{ and } \langle z, y\rangle \leq \alpha, \text{ for all } z \in \operatorname{cl} A.$$

This implies

$$h_{\operatorname{cl} A}(y) = h_A(y) \leq \alpha < \langle x, y\rangle,$$

a contradiction.

(c) The first part is obvious, the second follows from (b).

(d) If A is bounded, we have $A \subset B(r)$, for some $r > 0$. Then, (c) implies $h_A \leq h_{B(r)} = r\|\cdot\|$, hence $h_A < \infty$. Conversely, $h_A < \infty$ and Theorem 2.2.1 imply that h_A is continuous on $\mathbb{R}^n$. Therefore, h_A is bounded on S^{n-1}, i.e. $h_A \leq r = h_{B(r)}$ on S^{n-1}, for some $r > 0$. The positive homogeneity, proved in (a), implies that $h_A \leq h_{B(r)}$ on all of $\mathbb{R}^n$, hence (c) shows that $\operatorname{cl} A \subset B(r)$, i.e. A is bounded.

(e) For any $u \in \mathbb{R}^n$, we have

$$\begin{aligned} h_{\alpha A+\beta B}(u) &= \sup_{x \in \alpha A+\beta B} \langle x, u\rangle = \sup_{y \in A, z \in B} \langle \alpha y + \beta z, u\rangle = \sup_{y \in A} \langle \alpha y, u\rangle + \sup_{z \in B} \langle \beta z, u\rangle \\ &= \alpha h_A(u) + \beta h_B(u). \end{aligned}$$

(f) For any $u \in \mathbb{R}^n$, we have

$$\begin{aligned} h_{-A}(u) &= \sup_{x \in -A} \langle x, u\rangle = \sup_{y \in A} \langle -y, u\rangle \\ &= \sup_{y \in A} \langle y, -u\rangle = h_A(-u). \end{aligned}$$

(g) Since $A_i \subset A$, we have $h_{A_i} \leq h_A$ (from (c)), hence

$$\sup_{i \in I} h_{A_i} \leq h_A.$$

Conversely, any $y \in A$ has a representation

$$y = \alpha_1 y_{i_1} + \cdots + \alpha_k y_{i_k},$$

with $k \in \mathbb{N}$, $y_{i_j} \in A_{i_j}$, $\alpha_j \geq 0$, $\sum \alpha_j = 1$ and $i_j \in I$. Therefore, we get

$$\begin{aligned} h_A(u) = \sup_{y \in A} \langle y, u\rangle &= \sup_{y_{i_j} \in A_{i_j}, \alpha_j \geq 0, \sum \alpha_j = 1, i_j \in I, k \in \mathbb{N}} \langle \alpha_1 y_{i_1} + \cdots + \alpha_k y_{i_k}, u\rangle \\ &= \sup_{\alpha_j \geq 0, \sum \alpha_j = 1, i_j \in I, k \in \mathbb{N}} (\alpha_1 h_{A_{i_1}}(u) + \cdots + \alpha_k h_{A_{i_k}}(u)) \leq \sup_{i \in I} h_{A_i}(u). \end{aligned}$$

(h) Since $A \subset A_i$, we have $h_A \leq h_{A_i}$ (from (c)), for all $i \in I$. Using the inclusion of the epigraphs, the definition of cl and conv for functions and (a), we obtain

$$h_A \leq \mathrm{cl}\,\mathrm{conv}\,(h_{A_i})_{i\in I}.$$

On the other hand, Theorem 2.1.3 shows that

$$g := \mathrm{cl}\,\mathrm{conv}\,(h_{A_i})_{i\in I}$$

is the supremum of all affine functions below g. Since g is positively homogeneous, we can concentrate on linear functions. [In fact, if $\langle \cdot, y\rangle + \alpha \leq g$, then $\alpha \leq 0$ since $0 + \alpha \leq g(0) = 0$. For all $u \in \mathbb{R}^n$ and $\lambda > 0$, we have $\langle \lambda u, y\rangle + \alpha \leq g(\lambda u)$. Hence $\langle u, y\rangle + \alpha/\lambda \leq g(u)$, and therefore $\langle u, y\rangle \leq g(u)$. This shows that the given estimate can be replaced by the stronger estimate $\langle \cdot, y\rangle \leq g$.]

Therefore, assume $\langle \cdot, y\rangle \leq g$, $y \in \mathbb{R}^n$, is such a function. Then,

$$\langle \cdot, y\rangle \leq h_{A_i}, \quad \text{for all } i \in I.$$

(c) implies that $y \in A_i$, $i \in I$, hence $y \in \bigcap_{i\in I} A_i = A$. Therefore,

$$\langle \cdot, y\rangle \leq h_A,$$

from which we get

$$g = \mathrm{cl}\,\mathrm{conv}\,(h_{A_i})_{i\in I} \leq h_A.$$

(i) For $x \in \mathbb{R}^n$, we have

$$\delta_A^*(x) = \sup_{y\in\mathbb{R}^n} (\langle x, y\rangle - \delta_A(y)) = \sup_{y\in A} \langle x, y\rangle = h_A(x),$$

hence $\delta_A^* = h_A$. □

The following result is crucial for the later considerations.

Theorem 2.3.2. *Let $h : \mathbb{R}^n \to (-\infty, \infty]$ be positively homogeneous, closed and convex. Then there exists a unique nonempty, closed and convex set $A \subset \mathbb{R}^n$ such that*

$$h_A = h.$$

Proof. The positive homogeneity implies that $h(0) = 0$, hence h is proper.

We consider h^*. For $\alpha > 0$, we obtain from the positive homogeneity

$$h^*(x) = \sup_{y\in\mathbb{R}^n} (\langle x, y\rangle - h(y)) = \sup_{y\in\mathbb{R}^n} (\langle x, \alpha y\rangle - h(\alpha y))$$

$$= \alpha \sup_{y\in\mathbb{R}^n} (\langle x, y\rangle - h(y)) = \alpha h^*(x).$$

Therefore, h^* can only obtain the values 0 and ∞. We put $A := \operatorname{dom} h^*$. By Theorem 2.1.4(a), A is nonempty, closed and convex, and

$$h^* = \delta_A.$$

Theorem 2.3.1(i) implies

$$h^{**} = \delta_A^* = h_A.$$

By Theorem 2.1.4(b), we have $h^{**} = h$, hence $h_A = h$.

The uniqueness of A follows from Theorem 2.3.1(b). □

We mention without proof a couple of further properties of support functions, which are mostly simple consequences of the definition or the last two theorems. In the following remarks, A is always a nonempty closed convex subset of $\mathbb{R}^n$.

Remarks. (1) We have $A = \{x\}$, if and only if $h_A = \langle x, \cdot \rangle$.

(2) We have $h_{A+x} = h_A + \langle x, \cdot \rangle$.

(3) A is origin-symmetric (i.e. $A = -A$), if and only if h_A is *even*, i.e. $h_A(x) = h_A(-x)$, for all $x \in \mathbb{R}^n$.

(4) We have $0 \in A$, if and only if $h_A \geq 0$.

Let $A \subset \mathbb{R}^n$ be nonempty, closed and convex. For $u \in \mathbb{R}^n \setminus \{0\}$, we consider the sets

$$E(u) := \{x \in \mathbb{R}^n : \langle x, u \rangle = h_A(u)\}$$

and

$$A(u) := A \cap E(u) = \{x \in A : \langle x, u \rangle = h_A(u)\}.$$

If $h_A(u) = \infty$, both sets are empty. If $h_A(u) < \infty$, then $E(u)$ is a hyperplane, which bounds A, but need not be a supporting hyperplane (see the example in Section 1.4), namely if $A(u) = \emptyset$. If $A(u) \neq \emptyset$, then $E(u)$ is a supporting hyperplane of A (at each point $x \in A(u)$) and $A(u)$ is the corresponding support set. We discuss now the support function of $A(u)$. In order to simplify the considerations, we concentrate on the case, where A is compact (then $A(u)$ is nonempty and compact, for all $u \in S^{n-1}$).

Definition. A compact convex set $K \neq \emptyset$ is called a *convex body*. We denote by $\mathcal{K}^n$ the set of all convex bodies in $\mathbb{R}^n$.

Theorem 2.3.3. *Let $K \in \mathcal{K}^n$ and $u \in \mathbb{R}^n \setminus \{0\}$. Then,*

$$h_{K(u)}(x) = h'_K(u; x), \quad x \in \mathbb{R}^n,$$

i.e. the support function of $K(u)$ is given by the directional derivatives of h_K at the point u.

Proof. For $y \in K(u)$ and $v \in \mathbb{R}^n$, we have

$$\langle y, v \rangle \leq h_K(v),$$

since y belongs to K. In particular, for $v := u + tx, x \in \mathbb{R}^n, t > 0$, we thus get

$$\langle y, u\rangle + t\langle y, x\rangle \leq h_K(u + tx),$$

and hence

$$\langle y, x\rangle \leq \frac{h_K(u + tx) - h_K(u)}{t}$$

(because of $h_K(u) = \langle y, u\rangle$). For $t \searrow 0$, we obtain

$$\langle y, x\rangle \leq h'_K(u; x).$$

Since this holds for all $y \in K(u)$, we arrive at

$$h_{K(u)}(x) \leq h'_K(u; x). \tag{3.3}$$

Conversely, we obtain from the subadditivity of h_K

$$\frac{h_K(u + tx) - h_K(u)}{t} \leq \frac{h_K(tx)}{t} = h_K(x),$$

and thus

$$h'_K(u; x) \leq h_K(x).$$

This shows that the function $x \mapsto h'_K(u; x)$ is finite. As we have mentioned in the last section, it is also convex and positively homogeneous. Namely,

$$\begin{aligned} h'_K(u; x + z) &= \lim_{t \searrow 0} \frac{h_K(u + tx + tz) - h_K(u)}{t} \\ &\leq \lim_{t \searrow 0} \frac{h_K(\frac{u}{2} + tx) - h_K(\frac{u}{2})}{t} + \lim_{t \searrow 0} \frac{h_K(\frac{u}{2} + tz) - h_K(\frac{u}{2})}{t} \\ &\leq \lim_{t \searrow 0} \frac{h_K(u + 2tx) - h_K(u)}{2t} + \lim_{t \searrow 0} \frac{h_K(u + 2tz) - h_K(u)}{2t} \\ &= h'_K(u; x) + h'_K(u; z) \end{aligned}$$

and

$$h'_K(u; \alpha x) = \lim_{t \searrow 0} \frac{h_K(u + t\alpha x) - h_K(u)}{t} = \alpha h'_K(u; x),$$

for $x, z \in \mathbb{R}^n$ and $\alpha \geq 0$. By Theorem 2.3.2 (in connection with Theorem 2.3.1(d)), there exists a nonempty, compact convex set $L \subset \mathbb{R}^n$ with

$$h_L(x) = h'_K(u; x), \quad x \in \mathbb{R}^n.$$

For $y \in L$, we have

$$\langle y, x\rangle \leq h'_K(u; x) \leq h_K(x), \quad x \in \mathbb{R}^n,$$

hence $y \in K$. Furthermore,

$$\langle y, u\rangle \le h'_K(u; u) = h_K(u)$$

and

$$\langle y, -u\rangle \le h'_K(u; -u) = -h_K(u),$$

from which we obtain

$$\langle y, u\rangle = h_K(u),$$

and thus $y \in K \cap E(u) = K(u)$. It follows that $L \subset K(u)$, and therefore (again by Theorem 2.3.1)

$$h'_K(u; x) = h_L(x) \le h_{K(u)}(x). \tag{3.4}$$

Combining the inequalities (3.3) and (3.4), we obtain the assertion. □

Remark. As a consequence, we obtain that $K(u)$ consists of one point, if and only if $h'_K(u; \cdot)$ is linear. In view of Exercise 2.2.5 and Exercise 2.2.6, the latter is equivalent to the differentiability of h_K at u. If all the support sets $K(u), u \in S^{n-1}$, of a nonempty, compact convex set K consist of points, the boundary $\mathrm{bd}\, K$ does not contain any segments. Such sets K are called *strictly convex*. Hence, K is strictly convex, if and only if h_K is differentiable on $\mathbb{R}^n \setminus \{0\}$.

We finally consider the support functions of polytopes. We call a function h on $\mathbb{R}^n$ *piecewise linear*, if there are finitely many convex cones $A_1, \dots, A_m \subset \mathbb{R}^n$, such that $\mathbb{R}^n = \bigcup_{i=1}^m A_i$ and h is linear on A_i, $i = 1, \dots, m$.

Theorem 2.3.4. *Let $K \in \mathcal{K}^n$. Then K is a polytope, if and only if h_K is piecewise linear.*

Proof. The convex body K is a polytope, if and only if

$$K = \mathrm{conv}\,\{x_1, \dots, x_k\},$$

for some $x_1, \dots, x_k \in \mathbb{R}^n$. In view of Theorem 2.3.1, the latter is equivalent to

$$h_K = \max_{i=1,\dots,k} \langle x_i, \cdot\rangle,$$

which holds, if and only if h_K is piecewise linear.

To be more precise, if h_K has the above form, the convex cones A_i of linearity are given by

$$A_i := \{x \in \mathbb{R}^n : \max_{j=1,\dots,k} \langle x_j, x\rangle = \langle x_i, x\rangle\}, \quad i = 1, \dots, k.$$

Conversely, if h_K is linear on the cone A_i, we may assume that A_i is closed and has interior points. Then x_i is determined by

$$\langle x_i, \cdot\rangle = h_K$$

on A_i. The convexity of h_K implies that $h_K \ge \langle x_i, \cdot\rangle$ on $\mathbb{R}^n$. [In fact, let $z \in \mathrm{int}\, A_i$ and let $x \in \mathbb{R}^n \setminus \{z\}$. Then there are $y \in A_i$ and $\lambda \in (0, 1)$ such that $z = \lambda x + (1-\lambda)y$. Then

$$\langle x_i, z\rangle = h_K(z) = h_K(\lambda x + (1-\lambda)y) \le \lambda h_K(x) + (1-\lambda)h_K(y) = \lambda h_K(x) + (1-\lambda)\langle x_i, y\rangle,$$

and thus $\langle x_i, x\rangle \leq h_K(x)$ for all $x \in \mathbb{R}^n$.] Hence

$$h_K = \max_{i=1,\dots,k} \langle x_i, \cdot\rangle$$

follows. □

Exercises and problems

1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be positively homogeneous and twice continuously partially differentiable on $\mathbb{R}^n \setminus \{0\}$. Show that there are nonempty, compact convex sets $K, L \subset \mathbb{R}^n$ such that

$$f = h_K - h_L.$$

   Hint: Use Exercise 2.2.4(b).

2. Let $K \subset \mathbb{R}^n$ be compact and convex with $0 \in \operatorname{int} K$ and let K° be the polar of K (see Exercise 1.1.4). Show that

   (a) K° is compact and convex with $0 \in \operatorname{int} K^\circ$,

   (b) $K^{\circ\circ} := (K^\circ)^\circ = K$,

   (c) K is a polytope, if and only if K° is a polytope,

   (d) $h_K = d_{K^\circ}$.

Chapter 3

Convex bodies

3.1 The space of convex bodies

In the following, we mostly concentrate on *convex bodies* (nonempty compact convex sets) K in $\mathbb{R}^n$ and first discuss the space $\mathcal{K}^n$ of convex bodies. We emphasize that we do not require that a convex body has interior points; hence lower-dimensional bodies are included in $\mathcal{K}^n$. The set $\mathcal{K}^n$ is closed under addition,

$$K, L \in \mathcal{K}^n \implies K + L \in \mathcal{K}^n,$$

and multiplication with nonnegative scalars,

$$K \in \mathcal{K}^n, \alpha \geq 0 \implies \alpha K \in \mathcal{K}^n.$$

(In fact, we even have $\alpha K \in \mathcal{K}^n$, for all $\alpha \in \mathbb{R}$, since the reflection $-K$ of a convex body K is again a convex body.) Thus, $\mathcal{K}^n$ is a convex cone and the question arises, whether we can embed this cone into a suitable vector space. Since $(\mathcal{K}^n, +)$ is a (commutative) semi-group, the problem reduces to the question, whether this semi-group can be embedded into a group. A simple algebraic criterion (which is necessary and sufficient) is that the cancellation rule must be valid. Although this can be checked directly for convex bodies (see the exercises), we use now the support function for a direct embedding which has a number of additional advantages.

For this purpose, we consider the support function h_K of a convex body as a function on the unit sphere S^{n-1} (because of the positive homogeneity of h_K, the values on S^{n-1} determine h_K completely). Let $\mathbf{C}(S^{n-1})$ be the vector space of continuous functions on S^{n-1}. This is a Banach space with respect to the maximum norm

$$\|f\| := \max_{u \in S^{n-1}} |f(u)|, \quad f \in \mathbf{C}(S^{n-1}).$$

We call a function $f : S^{n-1} \to \mathbb{R}$ *convex*, if the homogeneous extension

$$\tilde{f} := \begin{cases} \|x\| f(\frac{x}{\|x\|}) & \text{for} \quad x \neq 0, \\ 0 & \phantom{\text{for}} \quad x = 0, \end{cases}$$

is convex on $\mathbb{R}^n$. Let $\mathcal{H}^n$ be the set of all convex functions on S^{n-1}. By Remark (3) (after Theorem 2.1.1) and Theorem 2.2.1, $\mathcal{H}^n$ is a convex cone in $\mathbf{C}(S^{n-1})$.

Theorem 3.1.1. *The mapping*

$$T : K \mapsto h_K$$

is (positively) linear on $\mathcal{K}^n$ and maps the convex cone $\mathcal{K}^n$ one-to-one onto the convex cone $\mathcal{H}^n$. Moreover, T is compatible with the inclusion order on $\mathcal{K}^n$ and the pointwise order $\leq$ on $\mathcal{H}^n$.

In particular, T embeds the (ordered) convex cone $\mathcal{K}^n$ into the (ordered) vector space $\mathbf{C}(S^{n-1})$.

Proof. The positive linearity of T follows from Theorem 2.3.1(e) and the injectivity from Theorem 2.3.1(b). The fact that $T(\mathcal{K}^n) = \mathcal{H}^n$ is a consequence of Theorem 2.3.2. The compatibility with respect to the orderings follows from Theorem 2.3.1(c). □

Remark. Positive linearity of T on the convex cone $\mathcal{K}^n$ means

$$T(\alpha K + \beta L) = \alpha T(K) + \beta T(L),$$

for $K, L \in \mathcal{K}^n$ and $\alpha, \beta \geq 0$. This linearity does not extend to negative α, β, in particular not to difference bodies $K - L = K + (-L)$. One reason is that the function $h_K - h_L$ is in general not convex, but even if it is, hence if

$$h_K - h_L = h_M,$$

for some $M \in \mathcal{K}^n$, the body M is in general different from the difference body $K - L$. We write $K \ominus L := M$ and call this body the *Minkowski difference* of K and L. Whereas the difference body $K - L$ exists for all $K, L \in \mathcal{K}^n$, the Minkowski difference $K \ominus L$ only exists in special cases, namely if K can be decomposed as $K = M + L$ (then $M = K \ominus L$).

With respect to the norm topology provided by the maximum norm in $\mathbf{C}(S^{n-1})$, the cone $\mathcal{H}^n$ is closed (see Exercise 6). Our next goal is to define a natural metric on $\mathcal{K}^n$, such that T becomes even an isometry (hence, we then have an isometric embedding of $\mathcal{K}^n$ into the Banach space $\mathbf{C}(S^{n-1})$).

Definition. For $K, L \in \mathcal{K}^n$, let

$$d(K, L) := \inf \{\varepsilon \geq 0 : K \subset L + B(\varepsilon), L \subset K + B(\varepsilon)\}.$$

It is easy to see that the infimum is attained, hence it is in fact a minimum.

Theorem 3.1.2. *For $K, L \in \mathcal{K}^n$, we have*

$$d(K, L) = \|h_K - h_L\|.$$

Therefore, d is a metric on $\mathcal{K}^n$ and fulfills

$$d(K + M, L + M) = d(K, L),$$

for all $K, L, M \in \mathcal{K}^n$.

Proof. From Theorem 2.3.1 we obtain

$$K \subset L + B(\varepsilon) \Leftrightarrow h_L \leq h_K + \varepsilon h_{B(1)}$$

and

$$L \subset K + B(\varepsilon) \Leftrightarrow h_K \leq h_L + \varepsilon h_{B(1)}.$$

Since $h_{B(1)} \equiv 1$ on S^{n-1}, this implies

$$K \subset L + B(\varepsilon), L \subset K + B(\varepsilon) \Leftrightarrow \|h_K - h_L\| \leq \varepsilon,$$

and the assertions follow. □

In an arbitrary metric space (X, d), the class $\mathcal{C}(X)$ of nonempty compact subsets of X can be supplied with the *Hausdorff metric* $\tilde{d}$, which is defined by

$$\tilde{d}(A, B) := \max\,(\max_{x \in A} d(x, B), \max_{y \in B} d(y, A)).$$

Here $A, B \in \mathcal{C}(X)$, and we have used the abbreviation

$$d(u, C) := \min_{v \in C} d(u, v), \quad u \in X, C \in \mathcal{C}(X),$$

(the minimal and maximal values exist due to the compactness of the sets and the continuity of the metric). We show now that, on $\mathcal{K}^n \subset \mathcal{C}(\mathbb{R}^n)$, the Hausdorff metric $\tilde{d}$ coincides with the metric d.

Theorem 3.1.3. *For $K, L \in \mathcal{K}^n$, we have*

$$d(K, L) = \tilde{d}(K, L).$$

Proof. We have

$$d(K, L) = \max\,(\inf\{\varepsilon \geq 0 : K \subset L + B(\varepsilon)\}, \inf\{\varepsilon \geq 0 : L \subset K + B(\varepsilon)\}).$$

Now

$$K \subset L + B(\varepsilon) \;\Leftrightarrow\; d(x, L) \leq \varepsilon, \quad \text{for all } x \in K,$$

$$\Leftrightarrow\; \max_{x \in K} d(x, L) \leq \varepsilon,$$

hence

$$\inf\{\varepsilon \geq 0 : K \subset L + B(\varepsilon)\} = \max_{x \in K} d(x, L),$$

which yields the assertion. □

We now come to an important topological property of the metric space $(\mathcal{K}^n, d)$: Every bounded subset $\mathcal{M} \subset \mathcal{K}^n$ is relative compact. This is a special property which holds also, for example, in the metric space $(\mathbb{R}^n, d)$, but it does not hold in general metric spaces.

In $\mathcal{K}^n$, a subset $\mathcal{M}$ is bounded, if there exists $c > 0$ such that

$$d(K, L) \leq c, \quad \text{for all } K, L \in \mathcal{M}.$$

This is equivalent to

$$K \subset B(c'), \quad \text{for all } K \in \mathcal{M},$$

for some constant $c' > 0$. Here, we can replace the ball $B(c')$ by any compact set, in particular by a cube $W \subset \mathbb{R}^n$. The subset $\mathcal{M}$ is relative compact, if every sequence $K_1, K_2, \dots$, with $K_k \in \mathcal{M}$, has a converging subsequence. Therefore, the mentioned topological property is a consequence of the following theorem.

Theorem 3.1.4 (BLASCHKE's Selection Theorem)**.** *Let $\mathcal{M} \subset \mathcal{K}^n$ be an infinite collection of convex bodies, all lying in a cube W. Then, there exists a sequence $K_1, K_2, \dots$, with $K_k \in \mathcal{M}$ (pairwise different), and a body $K_0 \in \mathcal{K}^n$ such that*

$$K_k \to K_0, \quad \textit{as } k \to \infty.$$

Proof. W.l.o.g. we assume that W is the unit cube.

For each $i \in \mathbb{N}$, we divide W into 2^{in} cubes of edge length $1/2^i$. For $K \in \mathcal{M}$, let $W_i(K)$ be the union of all cubes in the ith dissection, which intersect K. Since there are only finitely many different sets $W_i(K), K \in \mathcal{M}$, but infinitely many bodies $K \in \mathcal{M}$, we first get a sequence (in $\mathcal{M}$)

$$K_1^{(1)}, K_2^{(1)}, \dots$$

with

$$W_1(K_1^{(1)}) = W_1(K_2^{(1)}) = \cdots,$$

then a subsequence (of $K_1^{(1)}, K_2^{(1)}, \dots$)

$$K_1^{(2)}, K_2^{(2)}, \dots$$

with

$$W_2(K_1^{(2)}) = W_2(K_2^{(2)}) = \cdots,$$

and in general a subsequence

$$K_1^{(j)}, K_2^{(j)}, \dots$$

of $K_1^{(j-1)}, K_2^{(j-1)}, \dots$ with

$$W_j(K_1^{(j)}) = W_j(K_2^{(j)}) = \cdots,$$

for all $j \in \mathbb{N}$ ($j \geq 2$).

Since

$$\min_{y \in K_l^{(j)}} d(x,y) \le \frac{\sqrt{n}}{2^j},$$

for all $x \in K_k^{(j)}$, we have

$$d(K_k^{(j)}, K_l^{(j)}) \le \frac{\sqrt{n}}{2^j}, \quad \text{for all } k,l \in \mathbb{N}, \text{ and all } j.$$

By the subsequence property we deduce

$$d(K_k^{(j)}, K_l^{(i)}) \le \frac{\sqrt{n}}{2^i}, \quad \text{for all } k,l \in \mathbb{N}, \text{ and all } j \ge i.$$

In particular, if we choose the 'diagonal sequence' $K_k := K_k^{(k)}$, $k = 1, 2, \ldots$, then

$$d(K_k, K_l) \le \frac{\sqrt{n}}{2^l}, \quad \text{for all } k \ge l.$$

Hence $(K_k)_{k \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{M}$, that is, for each $\varepsilon > 0$ there exists $m \in \mathbb{N}$ such that

$$d(K_k, K_l) < \varepsilon, \quad \text{for all } k,l \ge m. \tag{1.1}$$

Let

$$\tilde{K}_k := \operatorname{cl}\operatorname{conv}(K_k \cup K_{k+1} \cup \cdots)$$

and

$$K_0 := \bigcap_{k=1}^{\infty} \tilde{K}_k.$$

We claim that

$$K_k \to K_0, \text{ as } k \to \infty, \quad \text{and} \quad K_0 \in \mathcal{K}^n.$$

First, by construction we have $\tilde{K}_k \in \mathcal{K}^n$ and $\tilde{K}_{k+1} \subset \tilde{K}_k$, $k = 1, 2, \ldots$. Therefore, $K_0 \neq \emptyset$ and hence $K_0 \in \mathcal{K}^n$.

For $\varepsilon > 0$, (1.1) implies

$$K_l \subset K_k + B(\varepsilon), \quad \text{for all } k,l \ge m,$$

therefore

$$\tilde{K}_{k'} \subset K_k + B(\varepsilon), \quad \text{for all } k,k' \ge m,$$

and thus

$$K_0 \subset K_k + B(\varepsilon), \quad \text{for all } k \ge m.$$

Conversely, for each $\varepsilon > 0$, there is $\bar{m} \in \mathbb{N}$ such that

$$\tilde{K}_k \subset K_0 + B(\varepsilon), \quad \text{for all } k \ge \bar{m}.$$

Namely, assume on the contrary that

$$\tilde{K}_k \not\subset K_0 + B(\varepsilon), \quad \text{for infinitely many } k.$$

Then

$$\tilde{K}_{k_i} \cap [W \setminus \operatorname{int}(K_0 + B(\varepsilon))] \neq \emptyset,$$

for a suitable sequence $k_1, k_2, \dots$. Since $\tilde{K}_{k_i}$ and $W \setminus \operatorname{int}(K_0 + B(\varepsilon))$ are compact, this would imply

$$\bigcap_{i=1}^{\infty} \big(\tilde{K}_{k_i} \cap [W \setminus \operatorname{int}(K_0 + B(\varepsilon))]\big) = K_0 \cap [W \setminus \operatorname{int}(K_0 + B(\varepsilon))] \neq \emptyset,$$

a contradiction.

Since $\tilde{K}_k \subset K_0 + B(\varepsilon)$ implies $K_k \subset K_0 + B(\varepsilon)$, we obtain

$$d(K_0, K_k) \leq \varepsilon, \quad \text{for all } k \geq \max(m, \bar{m}).$$

□

The topology on $\mathcal{K}^n$ given by the Hausdorff metric allows us to introduce and study geometric functionals on convex bodies by first defining them for a special subclass, for example the class $\mathcal{P}^n$ of polytopes. Such a program requires that the geometric functionals under consideration have a continuity or monotonicity property and also that the class $\mathcal{P}^n$ of polytopes is dense in $\mathcal{K}^n$. We now discuss the latter aspect; geometric functionals will be investigated in the next section.

Theorem 3.1.5. *Let $K \in \mathcal{K}^n$ and $\varepsilon > 0$.*

(a) *There exists a polytope $P \in \mathcal{P}^n$ with $P \subset K$ and $d(K, P) \leq \varepsilon$.*

(b) *There exists a polytope $P \in \mathcal{P}^n$ with $K \subset P$ and $d(K, P) \leq \varepsilon$.*

(c) *If $0 \in \operatorname{relint} K$, then there exists a polytope $P \in \mathcal{P}^n$ with $P \subset K \subset (1+\varepsilon)P$.*
There is even a polytope $\tilde{P} \in \mathcal{P}^n$ with $\tilde{P} \subset \operatorname{relint} K$ and $K \subset \operatorname{relint}((1+\varepsilon)\tilde{P})$.

Proof. (a) The family

$$\{x + \operatorname{int} B(\varepsilon) : x \in \operatorname{bd} K\}$$

is an open covering of the compact set $\operatorname{bd} K$, therefore there exist $x_1, \dots, x_m \in \operatorname{bd} K$ with

$$\operatorname{bd} K \subset \bigcup_{i=1}^{m} (x_i + \operatorname{int} B(\varepsilon)).$$

Let

$$P := \operatorname{conv}\{x_1, \dots, x_m\},$$

then

$$P \subset K \quad \text{and} \quad \operatorname{bd} K \subset P + B(\varepsilon).$$

The latter implies $K \subset P + B(\varepsilon)$ and therefore $d(K, P) \leq \varepsilon$.

(b) For each $u \in S^{n-1}$, there is a supporting hyperplane $E(u)$ of K (in direction u). Let $A(u)$ be the open half-space of $E(u)$ which fulfills $A(u) \cap K = \emptyset$ ($A(u)$ has the form $\{\langle \cdot, u \rangle > h_K(u)\}$). Then, the family

$$\{A(u) : u \in S^{n-1}\}$$

is an open covering of the compact set $\operatorname{bd}(K + B(\varepsilon))$, since every $y \in \operatorname{bd}(K + B(\varepsilon))$ fulfills $y \notin K$ and is therefore separated from K by a supporting hyperplane $E = E(u)$ of K. Therefore there exist $u_1, \ldots, u_m \in S^{n-1}$ with

$$\operatorname{bd}(K + B(\varepsilon)) \subset \bigcup_{i=1}^{m} A(u_i).$$

Let

$$P := \bigcap_{i=1}^{m} (\mathbb{R}^n \setminus A(u_i)),$$

then

$$K \subset P.$$

Since $\mathbb{R}^n \setminus P = \bigcup_{i=1}^{m} A(u_i)$, we also have

$$P \subset K + B(\varepsilon),$$

and therefore $d(K, P) \leq \varepsilon$.

(c) W.l.o.g. we may assume that $\dim K = n$, hence $0 \in \operatorname{int} K$. If we copy the proof of (b) with $B(\varepsilon) = \varepsilon B(1)$ replaced by εK, we obtain a polytope P' with

$$K \subset P' \subset (1 + \varepsilon)K.$$

The polytope $P := \frac{1}{1+\varepsilon} P'$ then fulfills $0 \in \operatorname{int} P$ and

$$P \subset K \subset (1 + \varepsilon)P.$$

In particular, we get a polytope $\bar{P}$ with $0 \in \operatorname{int} \bar{P}$ and

$$\bar{P} \subset K \subset (1 + \frac{\varepsilon}{2})\bar{P}.$$

We choose $\tilde{P} := \delta \bar{P}$ with $0 < \delta < 1$. Then

$$\tilde{P} \subset \operatorname{relint} \bar{P} \subset \operatorname{relint} K.$$

If δ is close to 1, such that $(1 + \frac{\varepsilon}{2})\frac{1}{\delta} < 1 + \varepsilon$, then

$$K \subset (1 + \frac{\varepsilon}{2})\frac{1}{\delta}\tilde{P} \subset \operatorname{relint}((1 + \varepsilon)\tilde{P}).$$

□

Remarks. (1) The theorem shows that $\mathrm{cl}\,\mathcal{P}^n = \mathcal{K}^n$. One can even show that the metric space $\mathcal{K}^n$ is separable, since there is a countable dense set $\tilde{\mathcal{P}}^n$ of polytopes. For this purpose, the above proofs have to be modified such that the polytopes involved have vertices with rational coordinates.

(2) In the proof of (a), the polytope P which was constructed has its vertices on $\mathrm{bd}\,K$. If we use the open covering $\{x + \mathrm{int}\,B(\varepsilon) : x \in \mathrm{relint}\,K\}$ of K instead, we obtain a polytope P with $d(K, P) \le \varepsilon$ and $P \subset \mathrm{relint}\,K$.

There is also a simultaneous proof of (b) and the first part of (c), which uses (a). Namely, assuming $\dim K = n$ and $0 \in \mathrm{int}\,K$, the body K contains a ball $B(\alpha), \alpha > 0$. For given $\varepsilon \in (0, 1)$, by (a) there is some $P \in \mathcal{P}^n$, $P \subset K$, such that $d(K, P) < \frac{\alpha\varepsilon}{2}$. Hence

$$h_P(u) \ge h_K(u) - \frac{\alpha\varepsilon}{2} \ge \alpha\left(1 - \frac{\varepsilon}{2}\right) > 0, \qquad u \in S^{n-1},$$

and therefore $\alpha(1 - \varepsilon/2)B^n \subset P$. This shows that

$$P \subset K \subset P + \frac{\alpha\varepsilon}{2}B^n \subset P + \frac{\alpha\varepsilon}{2}\frac{1}{\alpha(1-\varepsilon/2)}P = \left(1 + \frac{\varepsilon/2}{1-\varepsilon/2}\right)P \subset (1+\varepsilon)P.$$

Thus we obtain (c) and also get

$$\|h_{(1+\varepsilon)P} - h_K\| \le \varepsilon\|h_P\| \le \varepsilon\|h_K\|,$$

which implies (b).

Exercises and problems

1. Let $K, L, M \in \mathcal{K}^n$. Without using support functions, show that:
   (a) For $u \in S^{n-1}$, we have
   $$K(u) + M(u) = (K + M)(u).$$
   (b) If $K + L \subset M + L$, then $K \subset M$ (generalized cancellation rule).
2. Let $(K_i)_{i\in\mathbb{N}}$ be a sequence in $\mathcal{K}^n$ and $K \in \mathcal{K}^n$. Show that $K_i \to K$ (in the Hausdorff metric), if and only if the following two conditions are fulfilled:
   (a) Each $x \in K$ is a limit point of a suitable sequence $(x_i)_{i\in\mathbb{N}}$ with $x_i \in K_i$, for all $i \in \mathbb{N}$.
   (b) For each sequence $(x_i)_{i\in\mathbb{N}}$ with $x_i \in K_i$, for all $i \in \mathbb{N}$, every culmination point lies in K.
3. (a) Let $K, M \in \mathcal{K}^n$ be convex bodies, which cannot be separated by a hyperplane (i.e., there is no hyperplane $\{f = \alpha\}$ with $K \subset \{f \le \alpha\}$ and $M \subset \{f \ge \alpha\}$). Further, let $(K_i)_{i\in\mathbb{N}}$ and $(M_i)_{i\in\mathbb{N}}$ be sequences in $\mathcal{K}^n$. Show that
   $$K_i \to K,\ M_i \to M \Longrightarrow K_i \cap M_i \to K \cap M.$$

(b) Let $K \in \mathcal{K}^n$ be a convex body and $E \subset \mathbb{R}^n$ an affine subspace with $E \cap \operatorname{int} K \neq \emptyset$. Further, let $(K_i)_{i\in\mathbb{N}}$ be a sequence in $\mathcal{K}^n$. Show that

$$K_i \to K \Longrightarrow (E \cap K_i) \to E \cap K.$$

Hint: Use Exercise 2 above.

4. Let $K \subset \mathbb{R}^n$ be compact. Show that:

   (a) There is a unique ball K_a of smallest diameter with $K \subset K_a$ (*circumball*).

   (b) If $\operatorname{int} K \neq \emptyset$, then there exists a ball K_i of maximal diameter with $K_i \subset K$ (*inball*).

5. A body $K \in \mathcal{K}^n$ is *strictly convex*, if it does not contain any segments in the boundary.

   (a) Show that the set of all strictly convex bodies in $\mathbb{R}^n$ is a G_δ-set in $\mathcal{K}^n$, i.e. intersection of countably many open sets in $\mathcal{K}^n$.

   $*$ (b) Show that the set of all strictly convex bodies in $\mathbb{R}^n$ is dense in $\mathcal{K}^n$.

6. Let $(K_i)_{i\in\mathbb{N}}$ be a sequence in $\mathcal{K}^n$, for which the support functions $h_{K_i}(u)$ converge to the values $h(u)$ of a function $h : S^{n-1} \to \mathbb{R}$, for each $u \in S^{n-1}$. Show that h is the support function of a convex body and that $h_{K_i} \to h$ uniformly on S^{n-1}.

7. Let P be a convex polygon in $\mathbb{R}^2$ with $\operatorname{int} P \neq \emptyset$. Show that:

   (a) There is a polygon P_1 and a triangle (or a segment) Δ with $P = P_1 + \Delta$.

   (b) P has a representation $P = \Delta_1 + \cdots + \Delta_m$, with triangles (segments) Δ_j which are pairwise not homothetic.

   (c) P is a triangle, if and only if $m = 1$.

$*$ 8. A body $K \in \mathcal{K}^n, n \geq 2$, is *indecomposable*, if $K = M + L$ implies $M = \alpha K + x$ and $L = \beta K + y$, for some $\alpha, \beta \geq 0$ and $x, y \in \mathbb{R}^n$. Show that:

   (a) If $P \in \mathcal{K}^n$ is a polytope and all 2-faces of P are triangles, P is indecomposable.

   (b) For $n \geq 3$, the set of indecomposable convex bodies is a dense G_δ-set in $\mathcal{K}^n$.

9. Let $\mathcal{I}^n$ be the set of convex bodies $K \in \mathcal{K}^n$, which are strictly convex **and** indecomposable.

   (a) Show that $\mathcal{I}^n$ is dense in $\mathcal{K}^n$.

   P (b) Find one element of $\mathcal{I}^n$.

3.2 Volume and surface area

The volume of a convex body $K \in \mathcal{K}^n$ can be defined as the Lebesgue measure $\lambda_n(K)$ of K. However, the convexity of K implies that the volume also exists in an elementary sense and, moreover, that also the surface area of K exists. Therefore, we now introduce both notions in an elementary way, first for polytopes and then for arbitrary convex bodies by approximation.

Since we will use a recursive definition on the dimension n, we first remark that the support set $K(u)$, $u \in S^{n-1}$, of a convex body K lies in a hyperplane parallel to $u^\perp$. Therefore, the orthogonal projection $K(u)|u^\perp$ is a translate of $K(u)$, and we can consider $K(u)|u^\perp$ as a convex body in $\mathbb{R}^{n-1}$ (if we identify $u^\perp$ with $\mathbb{R}^{n-1}$). Assuming that the volume is already defined in $\mathbb{R}^{n-1}$, we then denote by $V^{(n-1)}(K(u)|u^\perp)$ the $(n-1)$-dimensional volume of this projection. In principle, the identification of $u^\perp$ with $\mathbb{R}^{n-1}$ requires that we have given an orthonormal basis in $u^\perp$. However, it will be apparent that the quantities we define depend only on the Euclidean metric in $u^\perp$, hence they are independent of the choice of a basis.

Definition. Let $P \in \mathcal{P}^n$ be a polytope.

For $n = 1$, hence $P = [a, b]$ with $a \leq b$, we define $V^{(1)}(P) := b - a$ and $F^{(1)}(P) := 2$.

For $n \geq 2$, let

$$V^{(n)}(P) := \begin{cases} \dfrac{1}{n} \sum\limits_{(*)} h_P(u) V^{(n-1)}(P(u)|u^\perp) & \text{if} \quad \dim P \geq n-1\,, \\ 0 & \phantom{\text{if}} \quad \dim P \leq n-2\,, \end{cases}$$

and

$$F^{(n)}(P) := \begin{cases} \sum\limits_{(*)} V^{(n-1)}(P(u)|u^\perp) & \text{if} \quad \dim P \geq n-1\,, \\ 0 & \phantom{\text{if}} \quad \dim P \leq n-2\,, \end{cases}$$

where the summation $(*)$ is over all $u \in S^{n-1}$, for which $P(u)$ is a facet of P. We shortly write $V(P)$ for $V^{(n)}(P)$ and call this the *volume* of P. Similarly, we write $F(P)$ instead of $F^{(n)}(P)$ and call this the *surface area* of P.

For $\dim P = n-1$, there are two support sets of P which are facets, namely $P = P(u_0)$ and $P = P(-u_0)$, where u_0 is a normal vector to P. Since then $V^{(n-1)}(P(u_0)|u_0^\perp) = V^{(n-1)}(P(-u_0)|u_0^\perp)$ and $h_P(u_0) = -h_P(-u_0)$, we obtain $V(P) = 0$, in coincidence with the Lebesgue measure of P. Also, in this case, $F(P) = 2V^{(n-1)}(P(u_0)|u_0^\perp)$. For $\dim P \leq n-2$, the polytope P does not have any facets, hence $V(P) = 0$ and $F(P) = 0$.

Proposition 3.2.1. *The volume V and surface area F of polytopes P, Q have the following properties:*

(1) $V(P) = \lambda_n(P)$,

(2) *V and F are invariant with respect to rigid motions,*

(3) $V(\alpha P) = \alpha^n V(P), F(\alpha P) = \alpha^{n-1}F(P)$, *for* $\alpha \geq 0$,

(4) $V(P) = 0$, *if and only if* $\dim P \leq n-1$,

(5) *if* $P \subset Q$, *then* $V(P) \leq V(Q)$ *and* $F(P) \leq F(Q)$.

Proof. (1) We proceed by induction on n. The result is clear for $n = 1$. Let $n \geq 2$. As we have already mentioned, $V(P) = 0 = \lambda_n(P)$ if $\dim P \leq n-1$. For $\dim P = n$, let $P(u_1), \dots, P(u_k)$ be the facets of P. Then, we have

$$V(P) = \frac{1}{n}\sum_{i=1}^{k} h_P(u_i)V^{(n-1)}(P(u_i)|u_i^\perp),$$

where, by the inductive assumption, $V^{(n-1)}(P(u_i)|u_i^\perp)$ equals the $(n-1)$-dimensional Lebesgue measure (in $u_i^\perp$) of $P(u_i)|u_i^\perp$. We assume w.l.o.g. that $h_P(u_1), \dots, h_P(u_m) \geq 0$ and $h_P(u_{m+1}), \dots, h_P(u_k) < 0$, and consider the pyramid-shaped polytopes $P_i := \operatorname{conv}(P(u_i) \cup \{0\})$, $i = 1, \dots k$. Then $V(P_i) = \frac{1}{n}h_P(u_i)V^{(n-1)}(P(u_i)|u_i^\perp)$, $i = 1, \dots, m$, and $V(P_i) = -\frac{1}{n}h_P(u_i)V^{(n-1)}(P(u_i)|u_i^\perp)$, $i = m+1, \dots, k$. Hence,

$$\begin{aligned} V(P) &= \sum_{i=1}^{m} V(P_i) - \sum_{i=m+1}^{k} V(P_i) \\ &= \sum_{i=1}^{m} \lambda_n(P_i) - \sum_{i=m+1}^{k} \lambda_n(P_i) \\ &= \lambda_n(P). \end{aligned}$$

Here, we have used that the Lebesgue measure of the pyramid P_i is $\frac{1}{n}$ times the content of the base (here $V^{(n-1)}(P(u_i)|u_i^\perp)$) times the height (here $h_P(u_i)$). Moreover the Lebesgue measure of the pyramid parts outside P cancel out, and the pyramid parts inside P yield a dissection of P (into sets with disjoint interior).

(2), (3), (4) and the first part of (5) follow now directly from (1) (and the corresponding properties of the Lebesgue measure). It remains to show $F(P) \leq F(Q)$, for polytopes $P \subset Q$. We may assume $\dim Q = n$. Again, we denote the facets of P by $P(u_1), \dots, P(u_k)$. We make use of the following elementary inequality (a generalization of the triangle inequality),

$$V^{(n-1)}(P(u_i)|u_i^\perp) \leq \sum_{j \neq i} V^{(n-1)}(P(u_j)|u_j^\perp), \qquad i = 1, \dots, k. \tag{2.2}$$

In order to motivate (2.2), we project $P(u_j), j \neq i$, orthogonally onto the hyperplane $u_i^\perp$. The projections then cover $P(u_i)|u_i^\perp$. Since the projection does not increase the $(n-1)$-dimensional Lebesgue measure, (2.2) follows. The estimate (2.2) implies that

$$F(Q \cap H) \leq F(Q),$$

for any closed half-space $H \subset \mathbb{R}^n$. Since $P \subset Q$ is a finite intersection of half-spaces, we obtain $F(P) \leq F(Q)$ by successive truncation. □

Remarks. (1) In the proof of (1), we could have avoided the occurrence of 'outside' pyramids by the following argument. If $0 \in \operatorname{int} P$, the pyramid dissection of P shows $V(P) = \lambda_n(P)$. For small enough $t \in \mathbb{R}^n$, we then have $-t \in \operatorname{int} P$ and the corresponding dissection w.r.t. t shows that $V(P+t) = V(P)$. We use

$$V(P) = \frac{1}{n} \sum_{i=1}^{k} h_P(u_i) \lambda_{n-1}(P(u_i)|u_i^\perp)$$

(which follows from the inductive assumption) and the same formula for $P + t$ and observe that

$$h_{P+t}(u_i) = h_P(u_i) + \langle t, u_i \rangle \qquad \text{and} \qquad \lambda_{n-1}((P+t)(u_i)|u_i^\perp) = \lambda_{n-1}(P(u_i)|u_i^\perp).$$

It follows that

$$\sum_{i=1}^{k} \langle t, u_i \rangle \lambda_{n-1}(P(u_i)) = 0.$$

Since this holds for all small enough t, it must hold for all $t \in \mathbb{R}^n$. Thus

$$\sum_{i=1}^{k} u_i \lambda_{n-1}(P(u_i)) = 0,$$

which yields $V(P+t) = V(P)$ for all $t \in \mathbb{R}^n$. Therefore, the assumption $0 \in \operatorname{int} P$ can be made w.l.o.g. and we obtain $V(P) = \lambda_n(P)$, in general.

(2) We can now simplify our formulas for the volume $V(P)$ and the surface area $F(P)$ of a polytope P. First, since we have shown that our elementarily defined volume equals the Lebesgue measure and is thus translation invariant, we do not need the orthogonal projection of the facets anymore. Second, since $V^{(n-1)}(P(u)) = 0$, for $\dim P(u) \leq n-2$, we can sum over all $u \in S^{n-1}$. If we write, in addition, v instead of $V^{(n-1)}$, we obtain

$$V(P) = \frac{1}{n} \sum_{u \in S^{n-1}} h_P(u) v(P(u))$$

and

$$F(P) = \sum_{u \in S^{n-1}} v(P(u)).$$

These are the formulas which we will use, in the following.

For a convex body $K \in \mathcal{K}^n$, we define

$$V_+(K) := \inf_{P \supset K} V(P), \quad V_-(K) := \sup_{P \subset K} V(P),$$

and

$$F_+(K) := \inf_{P \supset K} F(P), \quad F_-(K) := \sup_{P \subset K} F(P),$$

(here $P \in \mathcal{P}^n$).

Theorem 3.2.2 (and Definition)**.** *Let $K \in \mathcal{K}^n$.*
(a) *We have*

$$V_+(K) = V_-(K) =: V(K)$$

and

$$F_+(K) = F_-(K) =: F(K).$$

$V(K)$ *is called the* volume *and* $F(K)$ *is called the* surface area *of* K.
(b) *Volume and surface area have the following properties:*

(b1) $V(K) = \lambda_n(K)$,

(b2) V *and* F *are invariant with respect to rigid motions,*

(b3) $V(\alpha K) = \alpha^n V(K), F(\alpha K) = \alpha^{n-1} F(K)$, *for* $\alpha \geq 0$,

(b4) $V(K) = 0$, *if and only if* $\dim K \leq n - 1$,

(b5) *if* $K \subset L$, *then* $V(K) \leq V(L)$ *and* $F(K) \leq F(L)$,

(b6) $K \mapsto V(K)$ *is continuous.*

Proof. (a) We first remark that for a polytope P the monotonicity of V and F (which was proved in Proposition 3.2.1(5)) shows that $V^+(P) = V(P) = V^-(P)$ and $F^+(P) = F(P) = F^-(P)$. Hence, the new definition of $V(P)$ and $F(P)$ is consistent with the old one.

For an arbitrary body $K \in \mathcal{K}^n$, we get from Proposition 3.2.1(5)

$$V_-(K) \leq V_+(K) \qquad \text{and} \qquad F_-(K) \leq F_+(K),$$

and by Proposition 3.2.1(2), $V_-(K), V_+(K), F_-(K)$ and $F_+(K)$ are motion invariant. After a suitable translation, we may therefore assume $0 \in \operatorname{relint} K$. For $\varepsilon > 0$, we then use Theorem 3.1.5(c) and find a polytope P with $P \subset K \subset (1 + \varepsilon)P$. From Proposition 3.2.1(3), we get

$$V(P) \leq V_-(K) \leq V_+(K) \leq V((1 + \varepsilon)P) = (1 + \varepsilon)^n V(P)$$

and

$$F(P) \leq F_-(K) \leq F_+(K) \leq F((1 + \varepsilon)P) = (1 + \varepsilon)^{n-1} F(P).$$

For $\varepsilon \to 0$, this proves (a).

(b1) – (b5) follow now directly for bodies $K \in \mathcal{K}^n$ ((b1) by approximation with polytopes; (b2) – (b5) partially by approximation or from the corresponding properties of the Lebesgue measure).

It remains to prove (b6). Consider a convergent sequence $K_i \to K$, $K_i, K \in \mathcal{K}^n$. In view of (b2), we can assume $0 \in \operatorname{relint} K$. Using again Theorem 3.1.5(c), we find a polytope P with $P \subset \operatorname{relint} K$, $K \subset \operatorname{relint}(1+\varepsilon)P$. If $\dim K = n$, we can choose a ball $B(r), r > 0$, with $K + B(r) \subset (1+\varepsilon)P$ (choose $r = \min_{u \in S^{n-1}}(h_{(1+\varepsilon)P}(u) - h_K(u))$). Then $K_i \subset (1+\varepsilon)P$, for $i \geq i_0$. Analogously, we can choose a ball $B(r'), r' > 0$, with $P + B(r') \subset K \subset K_i + B(r')$, for $i \geq i_1$. This implies $P \subset K_i$ (see Exercise 1(b) of Section 3.1). For $i \geq \max(i_0, i_1)$, we therefore obtain

$$V(P) - V((1+\varepsilon)P) \leq V(K_i) - V(K) \leq V((1+\varepsilon)P) - V(P),$$

and hence

$$\begin{aligned} |V(K_i) - V(K)| &\leq (1+\varepsilon)^n V(P) - V(P) \\ &\leq [(1+\varepsilon)^n - 1]V(K) \to 0, \end{aligned}$$

as $\varepsilon \to 0$. If $\dim K = j \leq n-1$, hence $V(K) = 0$, we have

$$K \subset \operatorname{int}((1+\varepsilon)P + \varepsilon W),$$

where W is a cube, centred at 0, with edge length 1 and dimension $n-j$, lying in the orthogonal space $(\operatorname{aff} K)^\perp$. As above, we obtain $K_i \subset (1+\varepsilon)P + \varepsilon W$, for $i \geq i_0$. Since

$$V((1+\varepsilon)P + \varepsilon W) \leq \varepsilon^{n-j}(1+\varepsilon)^j C$$

(where we can choose the constant C to be the j-dimensional Lebesgue measure of K), this gives us $V(K_i) \to 0 = V(K)$, as $\varepsilon \to 0$. □

Remark. We shall see in the next section that the surface area F is also continuous.

Exercises and problems

1. A convex body $K \in \mathcal{K}^2$ is called *universal cover*, if for each $L \in \mathcal{K}^2$ with diameter ≤ 1 there is a rigid motion g_L with $L \subset g_L K$.

   (a) Show that there is a universal cover K_0 with minimal area.

P (b) Find the shape and the area of K_0.

3.3 Mixed volumes

There is another, commonly used definition of the surface area of a set $K \subset \mathbb{R}^n$, namely as the derivative of the volume functional of the outer parallel sets of K, i.e.

$$F(K) = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(V(K + B(\varepsilon)) - V(K)).$$

We will see now, that our notion of surface area of a convex body K fulfills also this limit relation. In fact, we will show that $V(K+B(\varepsilon))$ is a polynomial in ε (this is the famous STEINER *formula*) and thereby get a whole family of geometric functionals. We start with an even more general problem and investigate how the volume

$$V(\alpha_1 K_1 + \cdots + \alpha_m K_m),$$

for $K_i \in \mathcal{K}^n, \alpha_i > 0$, depends on the variables $\alpha_1, \ldots, \alpha_m$. This will lead us to a family of mixed functionals of convex bodies, the *mixed volumes*.

Again, we first consider the case of polytopes. Since the recursive representation of the volume of a polytope P was based on the support sets (facets) of P, we discuss how support sets behave under linear combinations.

Proposition 3.3.1. *Let $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m > 0$, let $P_1, \ldots, P_m \in \mathcal{P}^n$ be polytopes, and let $u, v \in S^{n-1}$. Then,*
(a) $(\alpha_1 P_1 + \cdots + \alpha_m P_m)(u) = \alpha_1 P_1(u) + \cdots + \alpha_m P_m(u)$,
(b) $\dim\,(\alpha_1 P_1 + \cdots + \alpha_m P_m)(u) = \dim\,(P_1 + \cdots + P_m)(u)$,
(c) *if* $(P_1 + \cdots + P_m)(u) \cap (P_1 + \cdots + P_m)(v) \neq \emptyset$*, then*

$$(P_1 + \cdots + P_m)(u) \cap (P_1 + \cdots + P_m)(v) = (P_1(u) \cap P_1(v)) + \cdots + (P_m(u) \cap P_m(v)).$$

Proof. (a) By Theorem 2.3.1 and Theorem 2.3.3, for all $x \in \mathbb{R}^n$ we have

$$\begin{aligned} h_{(\alpha_1 P_1 + \cdots + \alpha_m P_m)(u)}(x) &= h'_{\alpha_1 P_1 + \cdots + \alpha_m P_m}(u; x) \\ &= \alpha_1 h'_{P_1}(u; x) + \cdots + \alpha_m h'_{P_m}(u; x) \\ &= \alpha_1 h_{P_1(u)}(x) + \cdots + \alpha_m h_{P_m(u)}(x) \\ &= h_{\alpha_1 P_1(u) + \cdots + \alpha_m P_m(u)}(x). \end{aligned}$$

Theorem 2.3.1 now yields the assertion.

(b) Let $P := P_1 + \cdots + P_m$ and $\tilde{P} := \alpha_1 P_1 + \cdots + \alpha_m P_m$. W.l.o.g. we may assume $0 \in \operatorname{relint} P_i(u)$, $i = 1, \ldots, m$. By Exercise 1.3.3 (a) it follows that $0 \in \operatorname{relint} P(u)$. We put

$$\alpha := \min_{i=1,\ldots,m} \alpha_i, \quad \beta := \max_{i=1,\ldots,m} \alpha_i.$$

Then, $0 < \alpha < \beta$ and (in view of (a))

$$\alpha P(u) \subset \tilde{P}(u) \subset \beta P(u),$$

that is, $\dim P(u) = \dim \tilde{P}(u)$.

(c) Using the notation introduced above, we assume $P(u) \cap P(v) \neq \emptyset$. Consider $x \in P(u) \cap P(v)$. Since $x \in P$, it has a representation $x = x_1 + \cdots + x_m$, with $x_i \in P_i$. Because of

$$h_P(u) = \langle x, u \rangle = \sum_{i=1}^{m} \langle x_i, u \rangle \leq \sum_{i=1}^{m} h_{P_i}(u) = h_P(u),$$

we get that $\langle x_i, u \rangle = h_{P_i}(u)$ and thus $x_i \in P_i(u)$, for $i = 1, ..., m$. In the same way, we obtain $x_i \in P_i(v)$, $i = 1, ..., m$.

Conversely, it is clear that any $x \in (P_1(u) \cap P_1(v)) + \cdots + (P_m(u) \cap P_m(v))$ fulfills $x \in P_1(u) + \cdots + P_m(u) = P(u)$ and $x \in P_1(v) + \cdots + P_m(v) = P(v)$, by (a). $\square$

In the proof of an important symmetry property of mixed volumes, we also need the following lemma.

Lemma 3.3.2. *Let $K \in \mathcal{K}^n$, let $u, v \in S^{n-1}$ be linearly independent, and let $w = \lambda u + \mu v$ with some $\lambda \in \mathbb{R}$ and $\mu > 0$. Then $K(u) \cap K(v) \neq \emptyset$ implies that $K(u) \cap K(v) = K(u)(w)$.*

Proof. Let $z \in K(u) \cap K(v)$ and $w = \lambda u + \mu v$ with some $\lambda \in \mathbb{R}$ and $\mu > 0$. Then $z \in K(u)$, hence $\langle z, u \rangle = h_K(u) = h_{K(u)}(u)$ and

$$\begin{aligned} h_{K(u)}(-u) &= \max\{\langle x, -u \rangle : x \in K(u)\} = \max\{-\langle x, u \rangle : x \in K(u)\} \\ &= \max\{-h_{K(u)}(u) : x \in K(u)\} = -h_{K(u)}(u) = -\langle z, u \rangle = \langle z, -u \rangle. \end{aligned}$$

Therefore we have $\langle z, \lambda u \rangle = h_{K(u)}(\lambda u)$ for all $\lambda \in \mathbb{R}$. We deduce

$$\begin{aligned} \langle z, w \rangle &= \langle z, \lambda u \rangle + \langle z, \mu v \rangle = h_{K(u)}(\lambda u) + h_K(\mu v) \geq h_{K(u)}(\lambda u) + h_{K(u)}(\mu v) \\ &\geq h_{K(u)}(\lambda z + \mu v) = h_{K(u)}(w) \geq \langle z, w \rangle, \end{aligned}$$

which yields $z \in K(u)(w)$.

Now let $z \in K(u)(w)$. There is some $x_0 \in K(u) \cap K(v) \neq \emptyset$. Then $\langle x_0, u \rangle = h_K(u) = \langle z, u \rangle$, since $z \in K(u)$, and $\langle x_0, v \rangle = h_K(v)$. By the preceding argument, $x_0 \in K(u)(w)$, and therefore

$$\lambda \langle z, u \rangle + \mu \langle z, v \rangle = \langle z, w \rangle = \langle x_0, w \rangle = \lambda \langle x_0, u \rangle + \mu \langle x_0, v \rangle,$$

hence $\langle z, v \rangle = \langle x_0, v \rangle = h_K(v)$, i.e. $z \in K(v)$. Thus it follows that $z \in K(u) \cap K(v)$. $\square$

In analogy to the recursive definition of the volume of a polytope, we now define the mixed volume of polytopes. Again, we use projections of the support sets (faces) in order to make the definition rigorous. After we have shown translation invariance of the functionals, the corresponding formulas will become simpler.

For polytopes $P_1, ..., P_k \in \mathcal{P}_n$, let $N(P_1, ..., P_k)$ denote the set of all facet normals of the convex polytope $P_1 + \cdots + P_k$.

Definition. For polytopes $P_1, ..., P_n \in \mathcal{P}^n$, we define the *mixed volume* $V^{(n)}(P_1, ..., P_n)$ of $P_1, ..., P_n$ recursively:

$$V^{(1)}(P_1) := V(P_1) = h_{P_1}(1) + h_{P_1}(-1) \ (= b - a, \text{ if } P_1 = [a, b] \text{ with } a \leq b), \quad \text{for } n = 1,$$

$$V^{(n)}(P_1, ..., P_n) := \frac{1}{n} \sum_{u \in N(P_1, ..., P_{n-1})} h_{P_n}(u) V^{(n-1)}(P_1(u)|u^\perp, ..., P_{n-1}(u)|u^\perp), \text{ for } n \geq 2.$$

Theorem 3.3.3. *The mixed volume* $V^{(n)}(P_1, ..., P_n)$ *of polytopes* $P_1, ..., P_n \in \mathcal{P}^n$ *is symmetric in the indices* $1, ..., n$*, independent of individual translations of the polytopes* $P_1, ..., P_n$*, and for* $\dim(P_1 + \cdots + P_n) \leq n - 1$*, we have* $V^{(n)}(P_1, ..., P_n) = 0$.

Furthermore, for $m \in \mathbb{N}$*,* $P_1, ..., P_m \in \mathcal{P}^n$*, and* $\alpha_1, ..., \alpha_m \geq 0$*, we have*

$$V(\alpha_1 P_1 + \cdots + \alpha_m P_m) = \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} \alpha_{i_1} \cdots \alpha_{i_n} V^{(n)}(P_{i_1}, ..., P_{i_n}). \tag{3.3}$$

For the proof, it is convenient to extend the k-dimensional mixed volume $V^{(k)}(Q_1, ..., Q_k)$ (which is defined for polytopes $Q_1, ..., Q_k$ in a k-dimensional linear subspace $E \subset \mathbb{R}^d$) to polytopes $Q_1, ..., Q_k \in \mathcal{P}^n$, which fulfill $\dim(Q_1 + \cdots + Q_k) \leq k$, namely by

$$V^{(k)}(Q_1, ..., Q_k) := V^{(k)}(Q_1|E, ..., Q_k|E),$$

where E is a k-dimensional subspace parallel to $Q_1 + \cdots + Q_k$, $1 \leq k \leq n - 1$. The translation invariance and the dimensional condition, which we will prove, show that this extension is consistent (and independent of E in case $\dim(Q_1 + \cdots + Q_k) < k$). In the following inductive proof, we already make use of this extension in order to simplify the presentation. In particular, in the induction step, we use the mixed volume $V^{(n-1)}(P_1(u), ..., P_{n-1}(u))$.

In addition, we extend the mixed volume to the empty set, namely as $V^{(n)}(P_1, ..., P_n) := 0$, if one of the sets P_i is empty.

Proof. We use induction on the dimension n.

For $n = 1$, the polytopes P_i are intervals and the mixed volume equals the (one-dimensional) volume $V^{(1)}$ (the length of the intervals), which is linear

$$V^{(1)}(\alpha_1 P_1 + \cdots + \alpha_m P_m) = \sum_{i=1}^{m} \alpha_i V^{(1)}(P_i).$$

Hence, (3.3) holds as well as the remaining assertions.

Now we assume that the assertions of the theorem are true for all dimensions $\leq n-1$, and we consider dimension $n \geq 2$. We first discuss the dimensional statement. If $\dim(P_1 + \cdots + P_n) \leq n - 1$, then either $N(P_1, \ldots, P_{n-1}) = \emptyset$ or $N(P_1, \ldots, P_{n-1}) = \{-u, u\}$, where u is the normal on $\operatorname{aff}(P_1 + \cdots + P_n)$. In the first case, we have $V^{(n)}(P_1, \ldots, P_n) = 0$, by definition; in the second case, we have

$$\begin{aligned}
&V^{(n)}(P_1,\dots,P_n)\\
&= \frac{1}{n}(h_{P_n}(u)V^{(n-1)}(P_1(u),\dots,P_{n-1}(u)) + h_{P_n}(-u)V^{(n-1)}(P_1(-u),\dots,P_{n-1}(-u))\\
&= \frac{1}{n}(h_{P_n}(u)V^{(n-1)}(P_1(u),\dots,P_{n-1}(u)) - h_{P_n}(u)V^{(n-1)}(P_1(u),\dots,P_{n-1}(u))\\
&= 0.
\end{aligned}$$

Next, we prove (3.3). If $\alpha_i = 0$, for a certain index i, the corresponding summand $\alpha_i P_i$ on the left-hand side can be deleted, as well as all summands on the right-hand side which contain this particular index i. Therefore, it is sufficient to consider the case $\alpha_1 > 0, ..., \alpha_m > 0$. By the definition of volume and Proposition 3.3.1,

$$\begin{aligned}
V(\alpha_1 P_1 + \cdots + \alpha_m P_m) &= \frac{1}{n} \sum_{u\in N(P_1,\dots,P_m)} h_{\sum_{i=1}^m \alpha_i P_i}(u)\, v\left((\sum_{i=1}^m \alpha_i P_i)(u) \right)\\
&= \sum_{i_n=1}^m \alpha_{i_n} \frac{1}{n} \sum_{u\in N(P_1,\dots,P_m)} h_{P_{i_n}}(u)\, v\left(\sum_{i=1}^m \alpha_i (P_i(u)|u^\perp) \right).
\end{aligned}$$

The inductive assumption implies

$$v\left(\sum_{i=1}^m \alpha_i (P_i(u)|u^\perp) \right) = \sum_{i_1=1}^m \cdots \sum_{i_{n-1}=1}^m \alpha_{i_1}\cdots\alpha_{i_{n-1}}\, V^{(n-1)}(P_{i_1}(u), ..., P_{i_{n-1}}(u)).$$

Hence, we obtain

$$\begin{aligned}
&V(\alpha_1 P_1 + \cdots + \alpha_m P_m)\\
&= \sum_{i_1=1}^m \cdots \sum_{i_{n-1}=1}^m \sum_{i_n=1}^m \alpha_{i_1}\cdots\alpha_{i_{n-1}}\alpha_{i_n} \frac{1}{n} \sum_{u\in N(P_1,\dots,P_m)} h_{P_{i_n}}(u)\, V^{(n-1)}(P_{i_1}(u), ..., P_{i_{n-1}}(u))\\
&= \sum_{i_1=1}^m \cdots \sum_{i_n=1}^m \alpha_{i_1}\cdots\alpha_{i_n}\, V^{(n)}(P_{i_1}, ..., P_{i_n}).
\end{aligned}$$

Here, we have used that, for a given set of indices $\{i_1, ..., i_n\}$, the summation over $N(P_1, ..., P_m)$ can be replaced by the summation over $N(P_{i_1}, ..., P_{i_{n-1}})$. Namely, for $u \notin N(P_{i_1}, ..., P_{i_{n-1}})$ the support set $P_{i_1}(u) + \cdots + P_{i_{n-1}}(u) = (P_{i_1} + \cdots + P_{i_{n-1}})(u)$ has dimension $\leq n-2$ and hence $V^{(n-1)}(P_{i_1}(u), ..., P_{i_{n-1}}(u)) = 0$. We will use this fact also in the following parts of the proof.

We next show the symmetry. Since $V^{(n-1)}(P_1(u), ..., P_{n-1}(u))$ is symmetric (in the indices) by the inductive assumption, it suffices to show

$$V^{(n)}(P_1, ..., P_{n-2}, P_{n-1}, P_n) = V^{(n)}(P_1, ..., P_{n-2}, P_n, P_{n-1}).$$

Moreover, we may assume that $P := P_1 + \cdots + P_n$ fulfills $\dim P = n$. By definition,

$$V^{(n-1)}(P_1(u), ..., P_{n-1}(u))$$
$$= \frac{1}{n-1} \sum_{\tilde{v} \in \tilde{N}} h_{P_{n-1}(u)}(\tilde{v}) V^{(n-2)}((P_1(u))(\tilde{v}), ..., (P_{n-2}(u))(\tilde{v})),$$

where we have to sum over the set $\tilde{N}$ of facet normals of $P(u)$ (in $u^\perp$). Formally, we would have to work with the projections (the shifted support sets) $P_1(u)|u^\perp, ..., P_{n-1}(u)|u^\perp$, but here we make use of our extended definition of the $(n-2)$-dimensional mixed volume and of the fact that

$$h_{P_{n-1}(u)|u^\perp}(\tilde{v}) = h_{P_{n-1}(u)}(\tilde{v}),$$

for all $\tilde{v} \perp u$. The facets of $P(u)$ are $(n-2)$-dimensional faces of P, thus they arise (because of $\dim P = n$) as intersections $P(u) \cap P(v)$ of the facet $P(u)$ with another facet $P(v)$ of P. Since $\dim P = n$, the case $v = -u$ cannot occur. If $P(u) \cap P(v)$ is a $(n-2)$-face of P, hence a facet of $P(u)$, the corresponding normal (in $u^\perp$) is given by $\tilde{v} = \|v|u^\perp\|^{-1}(v|u^\perp)$, hence it is of the form $\tilde{v} = \lambda u + \mu v$ with some $\lambda \in \mathbb{R}$ and $\mu > 0$.

By Proposition 3.3.1(c),

$$P(u) \cap P(v) = (P_1(u) \cap P_1(v)) + \cdots + (P_n(u) \cap P_n(v));$$

in particular, $P_i(u) \cap P_i(v) \neq \emptyset$ for $i = 1, \ldots, n$. For a $(n-2)$-face $P(u) \cap P(v)$ of P, we therefore obtain by Lemma 3.3.2

$$(P_i(u))(\tilde{v}) = P_i(u) \cap P_i(v), \qquad i = 1, ..., n-2,$$

which implies

$$V^{(n-1)}(P_1(u), ..., P_{n-1}(u))$$
$$= \frac{1}{n-1} \sum_{\substack{v \in N(P_1,...,P_n),\\ P(u)\cap P(v)\neq\emptyset}} h_{P_{n-1}(u)}\left(\frac{v|u^\perp}{\|v|u^\perp\|}\right) V^{(n-2)}(P_1(u) \cap P_1(v), ..., P_{n-2}(u) \cap P_{n-2}(v)).$$

Here, we may sum again over all $v \in N(P_1, ..., P_n)$, with $P(u) \cap P(v) \neq \emptyset$, since for those v, for which $P(u) \cap P(v) \neq \emptyset$ is not an $(n-2)$-face of P, the mixed volume $V^{(n-2)}(P_1(u) \cap P_1(v), ..., P_{n-2}(u) \cap P_{n-2}(v))$ vanishes by the inductive hypothesis. Also, for $n = 2$, the mixed volume $V^{(n-2)}(P_1(u) \cap P_1(v), ..., P_{n-2}(u) \cap P_{n-2}(v))$ is defined to be 1.

Let $\gamma(u, v)$ denote the (outer) angle between u and v, then

$$\|v|u^\perp\| = \sin\gamma(u, v), \;\; \langle u, v\rangle = \cos\gamma(u, v),$$

and hence

$$\frac{v|u^\perp}{\|v|u^\perp\|} = \frac{1}{\sin\gamma(u, v)} v - \frac{1}{\tan\gamma(u, v)} u.$$

For $x \in P_{n-1}(u) \cap P_{n-1}(v)$, we have

$$\begin{aligned} h_{P_{n-1}(u)}(\tilde{v}) = \langle x, \tilde{v} \rangle &= \frac{1}{\sin \gamma(u,v)} \langle x, v \rangle - \frac{1}{\tan \gamma(u,v)} \langle x, u \rangle \\ &= \frac{1}{\sin \gamma(u,v)} h_{P_{n-1}}(v) - \frac{1}{\tan \gamma(u,v)} h_{P_{n-1}}(u). \end{aligned}$$

Hence, altogether we obtain

$$\begin{aligned} &V^{(n)}(P_1, ..., P_{n-2}, P_{n-1}, P_n) \\ &= \frac{1}{n} \sum_{u \in N(P_1,...,P_n)} h_{P_n}(u)\, V^{(n-1)}(P_1(u), ..., P_{n-1}(u)) \\ &= \frac{1}{n(n-1)} \sum_{u,v \in N(P_1,...,P_n), v \neq \pm u} \Big[\frac{1}{\sin \gamma(u,v)} h_{P_n}(u) h_{P_{n-1}}(v) \\ &\qquad - \frac{1}{\tan \gamma(u,v)} h_{P_n}(u) h_{P_{n-1}}(u) \Big] V^{(n-2)}(P_1(u) \cap P_1(v), ..., P_{n-2}(u) \cap P_{n-2}(v)) \\ &= V^{(n)}(P_1, ..., P_{n-2}, P_n, P_{n-1}), \end{aligned}$$

and the symmetry is proved.

For the remaining assertion, we put $m = n$ in (3.3). Since the left-hand side of (3.3) is invariant with respect to individual translations of the polytopes P_i, the same holds true for the coefficients of the polynomial on the right-hand side, in particular for the coefficient $V^{(n)}(P_1, ..., P_n)$. Here we need the symmetry of the coefficients and we make use of the fact that the coefficients of a polynomial in several variables are uniquely determined, if they are chosen to be symmetric. □

Remark. In the following, we use similar abbreviations as in the case of volume,

$$V(P_1, ..., P_n) := V^{(n)}(P_1, ..., P_n)$$

and

$$v(P_1(u), ..., P_{n-1}(u)) := V^{(n-1)}(P_1(u), ..., P_{n-1}(u)).$$

As a special case of the polynomial expansion of volumes, we obtain

$$V(P_1 + \cdots + P_m) = \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} V(P_{i_1}, \ldots, P_{i_n}).$$

The question arises, whether this expansion can be inverted.

Corollary 3.3.4 (Inversion Formula)**.** *For $P_1, ..., P_n \in \mathcal{P}^n$, we have*

$$V(P_1, ..., P_n) = \frac{1}{n!} \sum_{k=1}^{n} (-1)^{n+k} \sum_{1 \le r_1 < \cdots < r_k \le n} V(P_{r_1} + \cdots + P_{r_k}).$$

Proof. We denote the right-hand side by $f(P_1, ..., P_n)$, then formula $(*)$ in Theorem 3.3.3 implies that $f(\alpha_1 P_1, ..., \alpha_n P_n)$ is a homogeneous polynomial of degree n in the variables $\alpha_1 \ge 0, ..., \alpha_n \ge 0$ (and with symmetric coefficients). Replacing P_1 by $\{0\}$, we have

$$\begin{aligned}
&(-1)^{n+1} n! f(\{0\}, P_2, ..., P_n) \\
&= \sum_{2 \le r \le n} V(P_r) - \left[\sum_{2 \le r \le n} V(\{0\} + P_r) + \sum_{2 \le r < s \le n} V(P_r + P_s) \right] \\
&+ \left[\sum_{2 \le r < s \le n} V(\{0\} + P_r + P_s) + \sum_{2 \le r < s < t \le n} V(P_r + P_s + P_t) \right] \\
&- \cdots \\
&= 0,
\end{aligned}$$

which means that $f(\{0\}, \alpha_2 P_2, ..., \alpha_n P_n) = f(0 \cdot P_1, \alpha_2 P_2, ..., \alpha_n P_n)$ is the zero polynomial. Consequently, in the polynomial $f(\alpha_1 P_1, ..., \alpha_n P_n)$, only those coefficients can be non-vanishing which contain the index 1. Replacing 1 subsequently by $2, ..., n$, we obtain that only the coefficient of $\alpha_1 \cdots \alpha_n$ can be non-zero. This coefficient occurs only once in the representation of f, namely for $k = n$ with $(r_1, ..., r_n) = (1, ..., n)$. Therefore, by Theorem 3.3.3, this coefficient must coincide with $V(P_1, ..., P_n)$. □

Theorem 3.3.5. *For convex bodies $K_1, ..., K_n \in \mathcal{K}^n$ and arbitrary approximating sequences $(P_1^{(j)})_{j \in \mathbb{N}}, ..., (P_n^{(j)})_{j \in \mathbb{N}}$ of polytopes, such that $P_i^{(j)} \to K_i$, $i = 1, ..., n$, as $j \to \infty$, the limit*

$$V(K_1, ..., K_n) = \lim_{j \to \infty} V(P_1^{(j)}, ..., P_n^{(j)})$$

exists and is independent of the choice of the approximating sequences $(P_i^{(j)})_{j \in \mathbb{N}}$. The number $V(K_1, ..., K_n)$ is called the mixed volume *of $K_1, ..., K_n$. The mapping $V : (\mathcal{K}^n)^n \to \mathbb{R}$ defined by $(K_1, ..., K_n) \mapsto V(K_1, ..., K_n)$ is called* mixed volume.

In particular,

$$V(K_1, ..., K_n) = \frac{1}{n!} \sum_{k=1}^{n} (-1)^{n+k} \sum_{1 \le r_1 < \cdots < r_k \le n} V(K_{r_1} + \cdots + K_{r_k}). \tag{3.4}$$

and, for $m \in \mathbb{N}$, $K_1, ..., K_m \in \mathcal{K}^n$ and $\alpha_1, ..., \alpha_m \ge 0$,

$$V(\alpha_1 K_1 + \cdots + \alpha_m K_m) = \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} \alpha_{i_1} \cdots \alpha_{i_n} V(K_{i_1}, ..., K_{i_n}). \tag{3.5}$$

Furthermore, for all $K, L, K_1, \dots, K_n \in \mathcal{K}^n$*,*
(a) $V(K, ..., K) = V(K)$ *and* $nV(K, ..., K, B(1)) = F(K)$*.*
(b) V *is symmetric.*
(c) V *is multilinear, i.e.*

$$V(\alpha K + \beta L, K_2, ..., K_n) = \alpha V(K, K_2, ..., K_n) + \beta V(L, K_2, ..., K_n),$$

for all $\alpha, \beta \geq 0$*.*
(d) $V(K_1 + x_1, ..., K_n + x_n) = V(K_1, ..., K_n)$ *for all* $x_1, \dots, x_n \in \mathbb{R}^n$*.*
(e) $V(gK_1, ..., gK_n) = V(K_1, ..., K_n)$ *for all rigid motions* g*.*
(f) V *is continuous, i.e.*

$$V(K_1^{(j)}, ..., K_n^{(j)}) \to V(K_1, ..., K_n),$$

whenever $K_i^{(j)} \to K_i, i = 1, ..., n$*.*
(g) $V \geq 0$ *and* V *is monotone in each argument.*

Proof. The existence of the limit

$$V(K_1, ..., K_n) = \lim_{j\to\infty} V(P_1^{(j)}, ..., P_n^{(j)}),$$

the independence from the approximating sequences and formula (3.4) follow from Corollary 3.3.4 and the continuity of the addition of convex bodies and of the volume functional. Equation (3.5) is a consequence of (3.3).

(d), (e) and (f) follow now directly from (3.4).

(a) For polytopes the relation $V(K, ..., K) = V(K)$ follows by induction and for general bodies K by approximation with polytopes; alternatively, one can obtain it from Corollary 3.3.4 and (3.4). Concerning the relation $nV(K, ..., K, B(1)) = F(K)$, again we first discuss the case $K \in \mathcal{P}^n$. Let $(Q_j)_{j\in\mathbb{N}}$ be a sequence of polytopes with $Q_j \to B(1)$. Then,

$$nV(K, ..., K, Q_j) \to nV(K, ..., K, B(1))$$

and also

$$nV(K, ..., K, Q_j) = \sum_{u\in N(K)} h_{Q_j}(u)v(K(u))$$

$$\to \sum_{u\in N(K)} h_{B(1)}(u)v(K(u)) = \sum_{u\in N(K)} v(K(u)) = F(K).$$

For the generalization to arbitrary bodies K, we approximate K from inside and outside by polytopes and use (f); here only the monotonicity of the surface area is needed, not the continuity (which we have not proved yet).

(b) follows from the corresponding property for polytopes.

(c) is a consequence of $(*_2)$, if we apply it to the linear combination

$$\alpha_1(\alpha K + \beta L) + \alpha_2 K_2 + \cdots \alpha_m K_m = \alpha_1 \alpha K + \alpha_1 \beta L + \alpha_2 K_2 + \cdots \alpha_m K_m$$

twice (once as a combination of m bodies and once as a combination of $m+1$ bodies), and then compare the coefficients. Alternatively, if all bodies are polytopes, the assertion follows from the definition and the symmetry of mixed volumes together with the additivity of support functions. The general case is obtained by approximation.

(g) Again it is sufficient to prove this for polytopes. Then $V \geq 0$ follows by induction and the formula

$$V(P_1, ..., P_n) = \frac{1}{n} \sum_{u \in N(P_1,...,P_{n-1})} h_{P_n}(u) v(P_1(u), ..., P_{n-1}(u)),$$

where we may assume, in view of (d), that $0 \in \operatorname{relint} P_n$, hence $h_{P_n} \geq 0$. If $P_n \subset Q_n$, then $h_{P_n} \leq h_{Q_n}$, hence

$$V(P_1, ..., P_n) \leq V(P_1, ..., P_{n-1}, Q_n),$$

by the same formula and since the mixed volume is nonnegative. □

Remarks. (1) In addition to $V \geq 0$, one can show that $V(K_1, \dots, K_n) > 0$, if and only if there exist segments $s_1 \subset K_1, \dots, s_n \subset K_n$ with linearly independent directions (see the exercises below).

(2) Theorem 3.3.5 (a) and (f) now imply the continuity of the surface area F.

Now we consider the parallel body $K + B(\alpha)$, $\alpha \geq 0$, of a body $K \in \mathcal{K}^n$. With the choice $m = 2$, $\alpha_1 := 1, \alpha_2 := \alpha$ and $K_1 := K, K_2 := B(1)$, Theorem 3.3.5 implies

$$\begin{aligned} V(K + B(\alpha)) &= V(K + \alpha B(1)) = V(\alpha_1 K_1 + \alpha_2 K_2) \\ &= \sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} \alpha_{i_1} \cdots \alpha_{i_n} V(K_{i_1}, ..., K_{i_n}) \qquad (3.6) \\ &= \sum_{i=0}^{n} \alpha^i \binom{n}{i} V(\underbrace{K, \dots, K}_{n-i}, \underbrace{B(1), \dots, B(1)}_{i}). \end{aligned}$$

The coefficients in this particular polynomial expansion deserve special attention.

Definition. For $K \in \mathcal{K}^n$,

$$W_i(K) := V(\underbrace{K, \dots, K}_{n-i}, \underbrace{B(1), \dots, B(1)}_{i})$$

is called the i-th *quermassintegral* of K, $i = 0, \dots, n$, and

$$V_j(K) = V_j^{(n)}(K) := \frac{\binom{n}{j}}{\kappa_{n-j}} W_{n-j}(K) = \frac{\binom{n}{j}}{\kappa_{n-j}} V(\underbrace{K, \dots, K}_{j}, \underbrace{B(1), \dots, B(1)}_{n-j})$$

is called the j-th *intrinsic volume* of K, $j = 0, \dots, n$. Here, κ_k is the volume of the k-dimensional unit ball. Since we have extended the mixed volume to the empty set, we also define

$$W_i(\emptyset) := V_j(\emptyset) := 0, \qquad i, j = 0, \dots, n.$$

Formula (3.6) directly implies the following result.

Theorem 3.3.6 (STEINER formula)**.** *For $K \in \mathcal{K}^n$ and $\alpha \geq 0$, we have*

$$V(K + B(\alpha)) = \sum_{i=0}^{n} \alpha^i \binom{n}{i} W_i(K),$$

respectively

$$V(K + B(\alpha)) = \sum_{j=0}^{n} \alpha^{n-j} \kappa_{n-j} V_j(K).$$

Remarks. (1) In particular, we get

$$F(K) = nW_1(K) = \lim_{\alpha \searrow 0} \frac{1}{\alpha}(V(K + B(\alpha)) - V(K)),$$

hence the surface area is the "derivative" of the volume functional.
(2) As a generalization of the STEINER formula (3.6), one can show that

$$V_k(K + B(\alpha)) = \sum_{j=0}^{k} \alpha^{k-j} \binom{n-j}{n-k} \frac{\kappa_{n-j}}{\kappa_{n-k}} V_j(K),$$

for $k = 0, \dots, n-1$ (see the exercises).
(3) Here we deduced the Steiner formula as a special case of the polynomial expansion of the volume of a general Minkowski combination of convex bodies, that is via the introduction of mixed volumes. Of course, it is possible to follow a more direct approach by decomposing the outer parallel set of a convex polytope P by the inverse images under the projection map of the relative interiors of the faces of P. The result for a general convex body then follows again by approximation with polytopes.

The quermassintegrals are the classical notation used in most of the older books. The name quermassintegral will become clear in chapter 4 where we discuss some projection formulas. The intrinsic volumes follow the more modern terminology. Their advantages are that the index j of V_j corresponds to the degree of homogeneity,

$$V_j(\alpha K) = \alpha^j V_j(K), \qquad K \in \mathcal{K}^n, \alpha \geq 0,$$

and that they are independent of the surrounding dimension, i.e. for a body $K \in \mathcal{K}^n$ with $\dim K = k < n$, we have

$$V_j^{(n)}(K) = V_j^{(k)}(K), \qquad j = 0, \dots, k,$$

(see Exercise 5).

The intrinsic volumes are important geometric functionals of a convex body. First, by definition,

$$V_n(K) = V(K, \dots, K) = V(K)$$

is the volume of K. Then,

$$2V_{n-1}(K) = nV(K, \ldots, K, B(1)) = F(K)$$

is the surface area of K (such that, for a body K of dimension $n-1$, $V_{n-1}(K)$ is the $(n-1)$-dimensional content of K). At the other end, $V_1(K)$ is proportional to the *mean width* of K. Namely,

$$\frac{\kappa_{n-1}}{n} V_1(K) = V(K, B(1), \ldots, B(1)).$$

Approximating the unit ball by polytopes, one can show that

$$V(K, B(1), \ldots, B(1)) = \frac{1}{n} \int_{S^{n-1}} h_K(u) du,$$

where the integration is with respect to the spherical Lebesgue measure. A rigorous proof of this fact will be given in Section 3.5. Since $b_K(u) := h_K(u) + h_K(-u)$ gives the *width* of K in direction u (the distance between the two parallel supporting hyperplanes), we obtain

$$\frac{1}{n} \int_{S^{n-1}} h_K(u) du = \frac{1}{2n} \int_{S^{n-1}} b_K(u) du = \frac{\kappa_n}{2} \overline{B}(K),$$

where

$$\overline{B}(K) := \frac{1}{n\kappa_n} \int_{S^{n-1}} b_K(u) du$$

denotes the mean width. Hence

$$V_1(K) = \frac{n\kappa_n}{2\kappa_{n-1}} \overline{B}(K).$$

Finally,

$$V_0(K) = \frac{1}{\kappa_n} W_n(K) = \begin{cases} 1 & \text{if} \quad K \neq \emptyset\,, \\ 0 & \phantom{\text{if}} \quad K = \emptyset\,. \end{cases}$$

is the EULER-POINCARÉ *characteristic* of K. It plays an important role in integral geometry (see Chapter 4). The other intrinsic volumes $V_j(K), 1 < j < n-1$, have interpretations as integrals of curvature functions, if the boundary of K is smooth, e.g. $V_{n-2}(K)$ is proportional to the integral mean curvature of K.

Remark. From Theorem 3.3.5 we obtain the following additional properties of the intrinsic volumes V_j:

- $K \mapsto V_j(K)$ is continuous,
- V_j is motion invariant,
- $V_j \geq 0$ and V_j is monotone.

Later, in Section 4.3, we shall discuss a further property of V_j, namely the additivity. The intrinsic volume V_j is *additive* in the sense that

$$V_j(K \cup M) + V_j(K \cap M) = V_j(K) + V_j(M),$$

for all $K, M \in \mathcal{K}^n$ such that $K \cup M \in \mathcal{K}^n$.

Exercises and problems

1. (a) Let $s_1, \dots, s_n \in \mathcal{K}^n$ be segments of the form $s_i = [0, x_i],\ x_i \in \mathbb{R}^n$. Show that

$$n!\, V(s_1, \dots, s_n) = |\det(x_1, \dots, x_n)|.$$

   (b) For $K_1, \dots, K_n \in \mathcal{K}^n$, show that $V(K_1, \dots, K_n) > 0$ if and only if there exist segments $s_i \subset K_i, i = 1, \dots, n$, with linearly independent directions.

2. (a) For $K, M \in \mathcal{K}^2$ show the inequality

$$V(K, M) \le \frac{1}{8}\, F(K)\, F(M).$$

   Hint: Use Exercise 7 in Section 3.1.

   ∗ (b) Show that equality holds in the above inequality, if and only if K, M are orthogonal segments (or if one of the bodies is a point).

3.∗ (a) For $K \in \mathcal{K}^2$ show the inequality

$$V(K, -K) \le \frac{\sqrt{3}}{18}\, F^2(K).$$

   P (b) Show that equality holds in the above inequality, if and only if K is an equilateral triangle (or a point).

4. For $K, K' \in \mathcal{K}^n$, show that

$$\int_{D(K,K')} dx = \sum_{j=0}^{n} \binom{n}{j} V(\underbrace{K, \dots, K}_{n-j}, \underbrace{-K', \dots, -K'}_{j}),$$

   where $D(K, K') := \{z \in \mathbb{R}^n \,:\, K \cap (K' + z) \neq \emptyset\}$.

5. For $K \in \mathcal{K}^n$, show that the intrinsic volume $V_j(K) = V_j^{(n)}(K)$ is independent of the dimension n, i.e. if $\dim K = k < n$, then

$$V_j^{(k)}(K) = V_j^{(n)}(K), \quad \text{for } 0 \le j \le k.$$

6. Suppose $K \in \mathcal{K}^n$ and L is a q-dimensional linear subspace of $\mathbb{R}^n$, $q \in \{0, \dots, n-1\}$. Let B_L denote the unit ball in L.

   Show that:

   (a) $$V(K + \alpha B_L) = \sum_{j=0}^{q} \alpha^{q-j} \kappa_{q-j} \int_{L^\perp} V_j(K \cap (L + x))\, d\lambda_{n-q}(x), \quad \text{for all } \alpha \ge 0.$$

   (b) The $(n-q)$-dimensional volume of the projection $K \,|\, L^\perp$ fulfills

$$V_{n-q}(K | L^\perp) = \frac{\binom{n}{q}}{\kappa_q} V(\underbrace{K, \dots, K}_{n-q}, \underbrace{B_L, \dots, B_L}_{q}).$$

Hint for (a): Use FUBINI's theorem in $\mathbb{R}^n = L \times L^\perp$ for the left-hand side and apply Exercise 5.

7. For a convex body $K \in \mathcal{K}^n$ and $\alpha \geq 0$, prove the following STEINER formula for the intrinsic volumes:

$$V_k(K + B(\alpha)) = \sum_{j=0}^{k} \alpha^{k-j} \binom{n-j}{n-k} \frac{\kappa_{n-j}}{\kappa_{n-k}} V_j(K) \quad (0 \leq k \leq n-1).$$

P 8. Prove the following Theorem of HADWIGER:

Let $f : \mathcal{K}^n \to \mathbb{R}$ be additive, motion invariant and continuous (resp. monotone). Then, there are constants $\beta_j \in \mathbb{R}$ (resp. $\beta_j \geq 0$), such that

$$f = \sum_{j=0}^{n} \beta_j V_j.$$

3.4 The BRUNN-MINKOWSKI Theorem

The BRUNN-MINKOWSKI Theorem was one of the first main results on convex bodies (proved around 1890). It says that, for convex bodies $K, L \in \mathcal{K}^n$, the function

$$t \mapsto \sqrt[n]{V(tK + (1-t)L)}, \qquad t \in [0,1],$$

is concave. As consequences we will get inequalities for mixed volumes, in particular the celebrated isoperimetric inequality.

We first need an auxiliary result.

Lemma 3.4.1. *For $\alpha \in (0,1)$ and $r, s, t > 0$,*

$$\left(\frac{\alpha}{r} + \frac{1-\alpha}{s}\right)[\alpha r^t + (1-\alpha)s^t]^{\frac{1}{t}} \geq 1$$

with equality, if and only if $r = s$.

Proof. The function $x \mapsto \ln x$ is strictly concave, therefore we have

$$\begin{aligned}
&\ln\{(\frac{\alpha}{r} + \frac{1-\alpha}{s})[\alpha r^t + (1-\alpha)s^t]^{\frac{1}{t}}\} \\
&\quad = \frac{1}{t}\ln(\alpha r^t + (1-\alpha)s^t) + \ln(\frac{\alpha}{r} + \frac{1-\alpha}{s}) \\
&\quad \geq \frac{1}{t}(\alpha \ln r^t + (1-\alpha)\ln s^t) + \alpha \ln\frac{1}{r} + (1-\alpha)\ln\frac{1}{s} \\
&\quad = 0
\end{aligned}$$

with equality if and only if $r = s$ (the use of the logarithm is possible since its argument is always positive). The strict monotonicity of the logarithm now proves the result. □

The following important inequality is known as the Brunn-Minkowski inequality. It has numerous applications to and connections with geometry, analysis and probability theory.

Theorem 3.4.2 (BRUNN-MINKOWSKI)**.** *For convex bodies $K, L \in \mathcal{K}^n$ and $\alpha \in (0,1)$,*

$$\sqrt[n]{V(\alpha K + (1-\alpha)L)} \geq \alpha \sqrt[n]{V(K)} + (1-\alpha)\sqrt[n]{V(L)}$$

with equality, if and only if K and L lie in parallel hyperplanes or K and L are homothetic.

Remark. K and L are homothetic, if and only if $K = \alpha L + x$ or $L = \alpha K + x$, for some $x \in \mathbb{R}^n, \alpha \geq 0$. This includes the case of points, i.e. K and L are always homothetic, if K or L is a point.

Proof. We distinguish four cases.

Case 1: K and L lie in parallel hyperplanes. Then also $\alpha K + (1-\alpha)L$ lies in a hyperplane, and hence $V(K) = V(L) = 0$ and $V(\alpha K + (1-\alpha)L) = 0$.

Case 2: We have $\dim K \leq n-1$ and $\dim L \leq n-1$, but K and L do not lie in parallel hyperplanes, i.e. $\dim(K+L) = n$. Then $\dim(\alpha K + (1-\alpha)L) = n$, for all $\alpha \in (0,1)$, hence

$$\sqrt[n]{V(\alpha K + (1-\alpha)L)} > 0 = \alpha \sqrt[n]{V(K)} + (1-\alpha)\sqrt[n]{V(L)},$$

for all $\alpha \in (0,1)$.

Case 3: We have $\dim K \leq n-1$ and $\dim L = n$ (or vice versa). Then, for $x \in K$, we obtain

$$\alpha x + (1-\alpha)L \subset \alpha K + (1-\alpha)L,$$

and thus

$$(1-\alpha)^n V(L) = V(\alpha x + (1-\alpha)L) \leq V(\alpha K + (1-\alpha)L)$$

with equality, if and only if $K = \{x\}$.

Case 4: We have $\dim K = \dim L = n$. We may assume $V(K) = V(L) = 1$. Namely, for general K, L, we put

$$\overline{K} := \frac{1}{\sqrt[n]{V(K)}} K, \qquad \overline{L} := \frac{1}{\sqrt[n]{V(L)}} L$$

and

$$\overline{\alpha} := \frac{\alpha \sqrt[n]{V(K)}}{\alpha \sqrt[n]{V(K)} + (1-\alpha)\sqrt[n]{V(L)}}.$$

Then

$$\sqrt[n]{V(\overline{\alpha}\overline{K} + (1-\overline{\alpha})\overline{L})} \geq 1$$

implies the BRUNN-MINKOWSKI inequality, which we have to prove. Moreover, K and L are homothetic, if and only if $\overline{K}$ and $\overline{L}$ are homothetic.

Thus, we assume $V(K) = V(L) = 1$ and we have to show that

$$V(\alpha K + (1-\alpha)L) \geq 1$$

with equality if and only if K, L are translates of each other. Because the volume is translation invariant, we can make the additional assumption that K and L have their center of gravity at 0, where the center of gravity of an n-dimensional convex body M is the point $c \in \mathbb{R}^n$ fulfilling

$$\langle c, u\rangle = \frac{1}{V(M)} \int_M \langle x, u\rangle dx,$$

for all $u \in S^{n-1}$. The equality case then reduces to the claim that $K = L$.

We now prove the Brunn-Minkowski theorem by induction on n. For $n = 1$, the Brunn-Minkowski inequality follows from the linearity of the 1-dimensional volume and we even have equality which corresponds to the fact that in $\mathbb{R}^1$ any two convex bodies (compact intervals) are homothetic. Now assume $n \geq 2$ and the assertion of the Brunn-Minkowski theorem is true in dimension $n-1$. We choose a unit vector $u \in S^{n-1}$ and denote by

$$E_\eta := \{\langle \cdot, u\rangle = \eta\}, \qquad \eta \in \mathbb{R},$$

the hyperplane in direction u with (signed) distance η from the origin. The function

$$f : [-h_K(-u), h_K(u)] \to [0,1], \qquad \beta \mapsto V(K \cap \{\langle \cdot, u\rangle \leq \beta\}),$$

is strictly increasing and continuous. Since

$$V(K \cap \{\langle \cdot, u\rangle \leq \beta\}) = \int_{-h_K(-u)}^{\beta} v(K \cap E_\eta)\, d\eta$$

by Fubini's theorem and since $\eta \mapsto v(K \cap E_\eta)$ is continuous on $(-h_K(-u), h_K(u))$, the function f is differentiable on $(-h_K(-u), h_K(u))$ and $f'(\beta) = v(K \cap E_\beta)$. Since f is invertible, the inverse function $\beta : [0,1] \to [-h_K(-u), h_K(u)]$, which is also strictly increasing and continuous satisfies $\beta(0) = -h_K(-u)$, $\beta(1) = h_K(u)$ and

$$\beta'(\tau) = \frac{1}{f'(\beta(\tau))} = \frac{1}{v(K \cap E_{\beta(\tau)})}, \qquad \tau \in (0,1).$$

Analogously, for the body L we obtain a function $\gamma : [0,1] \to [-h_L(-u), h_L(u)]$ with

$$\gamma'(\tau) = \frac{1}{v(L \cap E_{\gamma(\tau)})}, \qquad \tau \in (0,1).$$

Because of

$$\alpha(K \cap E_{\beta(\tau)}) + (1-\alpha)(L \cap E_{\gamma(\tau)}) \subset (\alpha K + (1-\alpha)L) \cap E_{\alpha\beta(\tau)+(1-\alpha)\gamma(\tau)},$$

for $\alpha, \tau \in [0,1]$, we obtain from the inductive assumption

$$\begin{aligned}
&V(\alpha K + (1-\alpha)L) \\
&= \int_{-\infty}^{\infty} v((\alpha K + (1-\alpha)L) \cap E_\eta) d\eta \\
&= \int_0^1 v((\alpha K + (1-\alpha)L) \cap E_{\alpha\beta(\tau)+(1-\alpha)\gamma(\tau)})(\alpha\beta'(\tau) + (1-\alpha)\gamma'(\tau)) d\tau \\
&\geq \int_0^1 v\left(\alpha(K \cap E_{\beta(\tau)}) + (1-\alpha)(L \cap E_{\gamma(\tau)})\right)\left(\frac{\alpha}{v(K \cap E_{\beta(\tau)})} + \frac{1-\alpha}{v(L \cap E_{\gamma(\tau)})}\right) d\tau \\
&\geq \int_0^1 \left[\alpha \sqrt[n-1]{v(K \cap E_{\beta(\tau)})} + (1-\alpha) \sqrt[n-1]{v(L \cap E_{\gamma(\tau)})}\right]^{n-1} \\
&\qquad\qquad \times \left(\frac{\alpha}{v(K \cap E_{\beta(\tau)})} + \frac{1-\alpha}{v(L \cap E_{\gamma(\tau)})}\right) d\tau.
\end{aligned}$$

Choosing $r := v(K \cap E_{\beta(\tau)})$, $s := v(L \cap E_{\gamma(\tau)})$ and $t := \frac{1}{n-1}$, we obtain from Lemma 3.4.1 that the integrand is ≥ 1, which yields the required inequality.

Now assume

$$V(\alpha K + (1-\alpha)L) = 1.$$

Then we must have equality in our last estimation, which implies that the integrand equals 1, for all τ. Again by Lemma 3.4.1, this yields that

$$v(K \cap E_{\beta(\tau)}) = v(L \cap E_{\gamma(\tau)}), \qquad \text{for all } \tau \in [0,1].$$

Therefore $\beta' = \gamma'$, hence the function $\beta - \gamma$ is a constant. Because the center of gravity of K is at the origin, we obtain

$$0 = \int_K \langle x, u \rangle dx = \int_{\beta(0)}^{\beta(1)} \eta v(K \cap E_\eta) d\eta = \int_{\beta(0)}^{\beta(1)} \eta f'(\eta) d\eta = \int_0^1 \beta(\tau) d\tau,$$

where the change of variables $\eta = \beta(\tau)$ was used. In an analogous way,

$$0 = \int_0^1 \gamma(\tau) d\tau.$$

Consequently,

$$\int_0^1 (\beta(\tau) - \gamma(\tau)) d\tau = 0$$

and therefore $\beta = \gamma$. In particular, we obtain

$$h_K(u) = \beta(1) = \gamma(1) = h_L(u).$$

Since u was arbitrary, $V(\alpha K + (1-\alpha)L) = 1$ implies $h_K = h_L$, and hence $K = L$.

Conversely, it is clear that $K = L$ implies $V(\alpha K + (1-\alpha)L) = 1$. □

Remark. Theorem 3.4.2 implies that the function

$$f(t) := \sqrt[n]{V(tK + (1-t)L)}$$

is concave on $[0,1]$. Namely, let $x, y, \alpha \in [0,1]$, then

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= \sqrt[n]{V([\alpha x + (1-\alpha)y]K + [1 - \alpha x - (1-\alpha)y]L)} \\ &= \sqrt[n]{V(\alpha[xK + (1-x)L] + (1-\alpha)[yK + (1-y)L])} \\ &\geq \alpha \sqrt[n]{V(xK + (1-x)L)} + (1-\alpha) \sqrt[n]{V(yK + (1-y)L)} \\ &= \alpha f(x) + (1-\alpha) f(y). \end{aligned}$$

As a consequence of Theorem 3.4.2, we obtain an inequality for mixed volumes which was first proved by MINKOWSKI.

Theorem 3.4.3. *For $K, L \in \mathcal{K}^n$,*

$$V(K,\dots,K,L)^n \geq V(K)^{n-1}V(L)$$

with equality, if and only if $\dim K \leq n-2$ *or* K *and* L *lie in parallel hyperplanes or* K *and* L *are homothetic.*

Proof. For $\dim K \leq n-1$, the inequality holds since the right-hand side is zero. Moreover, we then have equality, if and only if either $\dim K \leq n-2$ or K and L lie in parallel hyperplanes (compare Exercise 3.3.1). Hence, we now assume $\dim K = n$.

By Theorem 3.4.2 (similarly to the preceding remark), it follows that the function

$$f(t) := V(K+tL)^{\frac{1}{n}}, \qquad t \in [0,1],$$

is concave. Therefore

$$f^+(0) \geq f(1) - f(0) = V(K+L)^{\frac{1}{n}} - V(K)^{\frac{1}{n}}.$$

Since

$$f^+(0) = \frac{1}{n}V(K)^{\frac{1}{n}-1} \cdot nV(K,...,K,L),$$

we arrive at

$$V(K)^{\frac{1}{n}-1} \cdot nV(K,...,K,L) \geq V(K+L)^{\frac{1}{n}} - V(K)^{\frac{1}{n}} \geq V(L)^{\frac{1}{n}},$$

where we used the Brunn-Minkowski inequality in the end (with $t = \frac{1}{2}$). This implies the assertion. Equality holds if and only if equality holds in the Brunn-Minkowski inequality, which yields that K and L are homothetic. □

Corollary 3.4.4 (Isoperimetric inequality)**.** *Let $K \in \mathcal{K}^n$ be a convex body of dimension n. Then,*

$$\left(\frac{F(K)}{F(B(1))}\right)^n \geq \left(\frac{V(K)}{V(B(1))}\right)^{n-1}.$$

Equality holds, if and only if K is a ball.

Proof. We put $L := B(1)$ in Theorem 3.4.3 and get

$$V(K,\dots,K,B(1))^n \geq V(K)^{n-1}V(B(1))$$

or, equivalently,

$$\frac{n^nV(K,\dots,K,B(1))^n}{n^nV(B(1),\dots,B(1),B(1))^n} \geq \frac{V(K)^{n-1}}{V(B(1))^{n-1}}.$$

□

The isoperimetric inequality states that, among all convex bodies of given volume (given surface area), the balls have the smallest surface area (the largest volume).

Using $V(B(1)) = \kappa_n$ and $F(B(1)) = n\kappa_n$, we can re-write the inequality in the form

$$V(K)^{n-1} \leq \frac{1}{n^n \kappa_n} F(K)^n.$$

For $n = 2$ and using the common terminology $A(K)$ for the area (the "volume" in $\mathbb{R}^2$) and $L(K)$ for the boundary length (the "surface area" in $\mathbb{R}^2$), we obtain

$$A(K) \leq \frac{1}{4\pi} L(K)^2,$$

and, for $n = 3$,

$$V(K)^2 \leq \frac{1}{36\pi} F(K)^3.$$

Exchanging K and $B(1)$ in the proof above yields a similar inequality for the mixed volume $V(B(1), \dots B(1), K)$, hence we obtain the following corollary for the mean width $\overline{B}(K)$.

Corollary 3.4.5. *Let $K \in \mathcal{K}^n$ be a convex body. Then,*

$$\left(\frac{\overline{B}(K)}{\overline{B}(B(1))} \right)^n \geq \frac{V(K)}{V(B(1))}.$$

Equality holds, if and only if K is a ball.

Remark. Since $\overline{B}(K)$ is not greater than the diameter of K, the corollary implies an inequality for the diameter.

Using Theorem 3.4.2 and the second derivatives, we obtain in a similar manner inequalities of quadratic type.

Theorem 3.4.6. *For $K, L \in \mathcal{K}^n$,*

$$V(K, \dots, K, L)^2 \geq V(K, \dots, K, L, L)V(K). \tag{4.7}$$

The *proof* is left as an exercise. The case of equality is not known completely. Equality holds for homothetic bodies, but there are also non-homothetic bodies (with interior points) for which equality holds.

Replacing K or L in (4.7) by the unit ball, we obtain more special inequalities, for example (in $\mathbb{R}^3$)

$$\pi \overline{B}(K)^2 \geq F(K)$$

or

$$F(K)^2 \geq 6\pi \overline{B}(K)V(K).$$

Exercises and problems

1. Give a proof of Theorem 3.4.6.

2. The *diameter* $\mathrm{diam}(K)$ of a convex body $K \in \mathcal{K}^n$ is defined as

$$\mathrm{diam}(K) := \sup\{\|x - y\| : x, y \in K\}.$$

   (a) Prove that
$$\overline{B}(K) \leq \mathrm{diam}(K) \leq \frac{n\kappa_n}{2\kappa_{n-1}} \cdot \overline{B}(K).$$

   (b) If there is equality in one of the two inequalities, what can be said about K?

3. Let $K \in \mathcal{K}^n$ be an n-dimensional convex body. The *difference body* $D(K)$ of K is defined as the centrally symmetric convex body $D(K) := \frac{1}{2}(K + (-K))$. Show that

   (a) $D(K)$ has the same width as K in every direction.

   (b) $V(D(K)) \geq V(K)$ with equality if and only if K is centrally symmetric.

3.5 Surface area measures

In Section 3.3, we have shown that, for polytopes $P_1, \dots, P_n \in \mathcal{P}^n$, the mixed volume fulfills the formula

$$V(P_1, \dots, P_{n-1}, P_n) = \frac{1}{n} \sum_{u \in S^{n-1}} h_{P_n}(u) v(P_1(u), \dots, P_{n-1}(u)).$$

Here, the summation extends over all unit vectors u for which $v(P_1(u), \dots, P_{n-1}(u)) > 0$, that is, over all facet normals of the polytope $P_1 + \dots + P_{n-1}$. By approximation (and using the continuity of mixed volumes and support functions), we therefore get the same formula for arbitrary bodies $K_n \in \mathcal{K}^n$,

$$V(P_1, \dots, P_{n-1}, K_n) = \frac{1}{n} \sum_{u \in S^{n-1}} h_{K_n}(u) v(P_1(u), \dots, P_{n-1}(u)). \tag{5.8}$$

We define

$$S(P_1, \dots, P_{n-1}, \cdot) := \sum_{u \in S^{n-1}} v(P_1(u), \dots, P_{n-1}(u)) \varepsilon_u, \tag{5.9}$$

where ε_u denotes the Dirac measure in $u \in S^{n-1}$,

$$\varepsilon_u(A) := \begin{cases} 1 & \text{if} \quad u \in A, \\ 0 & \quad u \notin A, \end{cases}$$

(here, A runs through all Borel sets in S^{n-1}). Then, $S(P_1, \dots, P_{n-1}, \cdot)$ is a finite Borel measure on the unit sphere S^{n-1}, which is called the *mixed surface area measure* of the polytopes $P_1, \dots, P_{n-1}$. Equation (5.8) is then equivalent to

$$V(P_1, \dots, P_{n-1}, K_n) = \frac{1}{n} \int_{S^{n-1}} h_{K_n}(u) dS(P_1, \dots, P_{n-1}, u). \tag{5.10}$$

Our next goal is to extend this integral representation to arbitrary convex bodies $K_1, \dots, K_{n-1}$ (and thus to define mixed surface area measures for general convex bodies).

We first need an auxiliary result.

Lemma 3.5.1. *For convex bodies $K_1, \dots, K_{n-1}, K_n, K_n' \in \mathcal{K}^n$, we have*

$$|V(K_1, \dots, K_{n-1}, K_n) - V(K_1, \dots, K_{n-1}, K_n')|$$

$$\leq \|h_{K_n} - h_{K_n'}\| V(K_1, \dots, K_{n-1}, B(1)).$$

Proof. First, let $K_1, \dots, K_{n-1}$ be polytopes. Since $h_{B(1)} \equiv 1$ (on S^{n-1}), we obtain from (5.8) that

$$\begin{aligned}
&|V(K_1, \dots, K_{n-1}, K_n) - V(K_1, \dots, K_{n-1}, K_n')| \\
&= \frac{1}{n} \left| \sum_{u \in S^{n-1}} (h_{K_n}(u) - h_{K_n'}(u)) v(K_1(u), \dots, K_{n-1}(u)) \right| \\
&\leq \frac{1}{n} \sum_{u \in S^{n-1}} |h_{K_n}(u) - h_{K_n'}(u)| v(K_1(u), \dots, K_{n-1}(u)) \\
&\leq \frac{1}{n} \sup_{v \in S^{n-1}} |h_{K_n}(v) - h_{K_n'}(v)| \sum_{u \in S^{n-1}} v(K_1(u), \dots, K_{n-1}(u)) \\
&= \frac{1}{n} \|h_{K_n} - h_{K_n'}\| \sum_{u \in S^{n-1}} h_{B(1)}(u) v(K_1(u), \dots, K_{n-1}(u)) \\
&= \|h_{K_n} - h_{K_n'}\| V(K_1, \dots, K_{n-1}, B(1)).
\end{aligned}$$

By Theorem 3.3.5 (continuity of the mixed volume), the inequality extends to arbitrary convex bodies. □

Now we can extend (5.10) to arbitrary convex bodies.

Theorem 3.5.2. *For $K_1, \dots, K_{n-1} \in \mathcal{K}^n$, there exists a uniquely determined finite Borel measure $S(K_1, \dots, K_{n-1}, \cdot)$ on S^{n-1} such that*

$$V(K_1, \dots, K_{n-1}, K) = \frac{1}{n} \int_{S^{n-1}} h_K(u) dS(K_1, \dots, K_{n-1}, u),$$

for all $K \in \mathcal{K}^n$.

Proof. We consider the Banach space $\mathbf{C}(S^{n-1})$ and the linear subspace $\mathbf{C}^2(S^{n-1})$ of twice continuously differentiable functions. Here, a function f on S^{n-1} is called twice continuously differentiable, if the homogeneous extension $\tilde{f}$ of f,

$$\tilde{f}(x) := \begin{cases} \|x\| f(\frac{x}{\|x\|}) & \text{if} \quad x \in \mathbb{R}^n \setminus \{0\}, \\ 0 & \qquad\;\; x = 0, \end{cases}$$

is twice continuously differentiable on $\mathbb{R}^n \setminus \{0\}$. From analysis we use the fact that the subspace $\mathbf{C}^2(S^{n-1})$ is dense in $\mathbf{C}(S^{n-1})$, that is, for each $f \in \mathbf{C}(S^{n-1})$ there is a sequence of functions $f_i \in \mathbf{C}^2(S^{n-1})$ with $f_i \to f$ in the maximum norm, as $i \to \infty$ (this can be proved either by a convolution argument or by using a result of STONE-WEIERSTRASS type).

Further, we consider the set $\mathcal{L}^n$ of all functions $f \in \mathbf{C}(S^{n-1})$ which have a representation $f = h_K - h_{K'}$ with convex bodies $K, K' \in \mathcal{K}^n$. Obviously, $\mathcal{L}^n$ is also a linear subspace. Exercise 3.2.1 shows that $\mathbf{C}^2(S^{n-1}) \subset \mathcal{L}^n$, therefore $\mathcal{L}^n$ is dense in $\mathbf{C}(S^{n-1})$.

We now define a functional $T_{K_1,\dots,K_{n-1}}$ on $\mathcal{L}^n$ by

$$T_{K_1,\dots,K_{n-1}}(f) := nV(K_1,\dots,K_{n-1},K) - nV(K_1,\dots,K_{n-1},K'),$$

where $f = h_K - h_{K'}$. This definition is actually independent of the particular representation of f. Namely, if $f = h_K - h_{K'} = h_L - h_{L'}$, then $K + L' = K' + L$ and hence

$$\begin{aligned} &V(K_1,\dots,K_{n-1},K) + V(K_1,\dots,K_{n-1},L') \\ &\qquad = V(K_1,\dots,K_{n-1},K') + V(K_1,\dots,K_{n-1},L), \end{aligned}$$

by the multilinearity of mixed volumes. This yields

$$\begin{aligned} &nV(K_1,\dots,K_{n-1},K) - nV(K_1,\dots,K_{n-1},K') \\ &\qquad = nV(K_1,\dots,K_{n-1},L) - nV(K_1,\dots,K_{n-1},L'). \end{aligned}$$

The argument just given also shows that $T_{K_1,\dots,K_{n-1}}$ is linear. Moreover, $T_{K_1,\dots,K_{n-1}}$ is a positive functional since $f = h_K - h_{K'} \geq 0$ implies $K \supset K'$. Hence

$$V(K_1,\dots,K_{n-1},K) \geq V(K_1,\dots,K_{n-1},K')$$

and therefore $T_{K_1,\dots,K_{n-1}}(f) \geq 0$. Finally, $T_{K_1,\dots,K_{n-1}}$ is continuous (with respect to the maximum norm), since Lemma 3.5.1 shows that

$$|T_{K_1,\dots,K_{n-1}}(f)| \leq c(K_1,\dots,K_{n-1})\|f\|$$

with $c(K_1,\dots,K_{n-1}) := nV(K_1,\dots,K_{n-1},B(1))$.

Since $\mathcal{L}^n$ is dense in $\mathbf{C}(S^{n-1})$, the inequality just proven (or alternatively, the theorem of HAHN-BANACH) implies that there is a unique continuous extension of $T_{K_1,\dots,K_{n-1}}$ to a positive linear functional on $\mathbf{C}(S^{n-1})$. The RIESZ representation theorem then shows that

$$T_{K_1,\dots,K_{n-1}}(f) = \int_{S^{n-1}} f(u)dS(K_1,\dots,K_{n-1},u),$$

for $f \in \mathbf{C}(S^{n-1})$, with a finite (nonnegative) Borel measure $S(K_1,\dots,K_{n-1},\cdot)$ on S^{n-1}, which is uniquely determined by $T_{K_1,\dots,K_{n-1}}$. The existence assertion of the theorem now follows, if we put $f = h_K$.

For the uniqueness part, let μ, μ' be two Borel measures on S^{n-1}, depending on $K_1,\dots,K_{n-1}$, such that

$$\int_{S^{n-1}} h_K(u)d\mu(u) = \int_{S^{n-1}} h_K(u)d\mu'(u)$$

for all $K \in \mathcal{K}^n$. By linearity, we get

$$\int_{S^{n-1}} f(u)d\mu(u) = \int_{S^{n-1}} f(u)d\mu'(u)$$

first for all $f \in \mathcal{L}^n$, and then for all $f \in \mathbf{C}(S^{n-1})$. The uniqueness assertion in the Riesz representation theorem then implies that $\mu = \mu'$. □

Definition. The measure $S(K_1, \dots, K_{n-1}, \cdot)$ is called the *mixed surface area measure* of the bodies $K_1, \dots, K_{n-1}$. In particular,

$$S_j(K, \cdot) := S(\underbrace{K, \dots, K}_{j}, \underbrace{B(1), \dots, B(1)}_{n-1-j}, \cdot)$$

is called the jth order *surface area measure* of K, $j = 0, \dots, n-1$.

Remarks. (1) For polytopes $K_1, \dots, K_{n-1}$, the mixed surface area measure $S(K_1, \dots, K_{n-1}, \cdot)$ equals the measure defined in (5.9).

(2) All surface area measures have centroid 0. Namely, since

$$V(K_1, \dots, K_{n-1}, \{x\}) = 0,$$

we have

$$\int_{S^{n-1}} \langle x, u \rangle dS(K_1, \dots, K_{n-1}, u) = 0,$$

for all $x \in \mathbb{R}^n$.

(3) We have

$$\begin{aligned} S_j(K, S^{n-1}) &= nV(\underbrace{K, \dots, K}_{j}, \underbrace{B(1), \dots, B(1)}_{n-j}) \\ &= \frac{n\kappa_{n-j}}{\binom{n}{j}} V_j(K), \end{aligned}$$

in particular

$$S_{n-1}(K, S^{n-1}) = 2V_{n-1}(K) = F(K),$$

which explains the name surface area measure.

(4) The measure $S_0(K, \cdot) = S(B(1), \dots, B(1), \cdot) = S_j(B(1), \cdot)$ (for $j = 0, \dots, n-1$ and $K \in \mathcal{K}^n$) equals the spherical Lebesgue measure ω_{n-1} (this follows from part (d) of the following theorem), hence we obtain the equation

$$V(K, B(1), \dots, B(1)) = \frac{1}{n} \int_{S^{n-1}} h_K(u) du,$$

which we used already at the end of Section 3.3.

Further properties of mixed surface area measures follow, if we combine Theorem 3.5.2 with Theorem 3.3.5. In order to formulate a continuity result, we make use of the weak convergence of measures on S^{n-1} (since S^{n-1} is compact, weak and vague convergence are the same). A sequence of finite measures $\mu_i, i = 1, 2, \dots,$ on S^{n-1} is said to *converge weakly* to a finite measure μ on S^{n-1}, if and only if

$$\int_{S^{n-1}} f(u) d\mu_i(u) \to \int_{S^{n-1}} f(u) d\mu(u), \qquad \text{as } i \to \infty,$$

for all $f \in \mathbf{C}(S^{n-1})$.

Theorem 3.5.3. *The mapping $S : (K_1, \dots, K_{n-1}) \mapsto S(K_1, \dots, K_{n-1}, \cdot)$ has the following properties:*
(a) *S is symmetric, i.e.*

$$S(K_1, \dots, K_{n-1}, \cdot) = S(K_{\pi(1)}, \dots, K_{\pi(n-1)}, \cdot),$$

for all $K_1, ..., K_{n-1} \in \mathcal{K}^n$ and all permutations π of $1, \dots, n-1$.
(b) *S is multilinear, i.e.*

$$S(\alpha K + \beta L, K_2, ..., K_{n-1}, \cdot) = \alpha S(K, K_2, ..., K_{n-1}, \cdot) + \beta S(L, K_2, ..., K_{n-1}, \cdot),$$

for all $\alpha, \beta \geq 0, K, L, K_2, ..., K_{n-1} \in \mathcal{K}^n$.
(c) *S is translation invariant, i.e.*

$$S(K_1 + x_1, ..., K_{n-1} + x_{n-1}, \cdot) = S(K_1, ..., K_{n-1}, \cdot),$$

for all $K_i \in \mathcal{K}^n$ and all $x_i \in \mathbb{R}^n$.
(d) *S is rotation covariant, i.e.*

$$S(\vartheta K_1, ..., \vartheta K_{n-1}, \vartheta A) = S(K_1, ..., K_{n-1}, A),$$

for all $K_i \in \mathcal{K}^n$, all Borel sets $A \subset S^{n-1}$, and all rotations ϑ.
(e) *S is continuous, i.e.*

$$S(K_1^{(m)}, ..., K_{n-1}^{(m)}, \cdot) \to S(K_1, ..., K_{n-1}, \cdot)$$

weakly, as $m \to \infty$, provided $K_i^{(m)} \to K_i, i = 1, ..., n-1$.

Proof. (a), (b) and (c) follow directly from the integral representation and the uniqueness in Theorem 3.5.2 together with the corresponding properties of mixed volumes in Theorem 3.3.5.

(d) If $\rho \circ \mu$ denotes the image of a measure μ on S^{n-1} under the rotation ρ, then

$$\begin{aligned}
&\int_{S^{n-1}} h_{K_n}(u) d[\vartheta^{-1} \circ S(\vartheta K_1, ..., \vartheta K_{n-1}, \cdot)](u) \\
&\quad = \int_{S^{n-1}} h_{K_n}(\vartheta^{-1} u) dS(\vartheta K_1, ..., \vartheta K_{n-1}, u) \\
&\quad = \int_{S^{n-1}} h_{\vartheta K_n}(u) dS(\vartheta K_1, ..., \vartheta K_{n-1}, u) \\
&\quad = nV(\vartheta K_1, ..., \vartheta K_{n-1}, \vartheta K_n) \\
&\quad = nV(K_1, ..., K_{n-1}, K_n) \\
&\quad = \int_{S^{n-1}} h_{K_n}(u) dS(K_1, ..., K_{n-1}, u),
\end{aligned}$$

where $K_n \in \mathcal{K}^n$ is arbitrary. The assertion now follows from the uniquess part of Theorem 3.5.2.

(e) For $\varepsilon > 0$ and $f \in \mathbf{C}(S^{n-1})$, choose $K, L \in \mathcal{K}^n$ with

$$\|f - (h_K - h_L)\| \leq \varepsilon$$

and then m_0 such that $K_i^{(m)} \subset K_i + B(1)$, $i = 1, ..., n-1$, and

$$|V(K_1^{(m)}, \dots, K_{n-1}^{(m)}, K) - V(K_1, \dots, K_{n-1}, K)| \leq \varepsilon,$$

as well as

$$|V(K_1^{(m)}, \dots, K_{n-1}^{(m)}, L) - V(K_1, \dots, K_{n-1}, L)| \leq \varepsilon,$$

for all $m \geq m_0$. Then,

$$\begin{aligned}
&\left| \int_{S^{n-1}} f(u) dS(K_1^{(m)}, \dots, K_{n-1}^{(m)}, u) - \int_{S^{n-1}} f(u) dS(K_1, \dots, K_{n-1}, u) \right| \\
&\leq \left| \int_{S^{n-1}} (f - (h_K - h_L))(u) dS(K_1^{(m)}, \dots, K_{n-1}^{(m)}, u) \right| \\
&\quad + \left| \int_{S^{n-1}} (h_K - h_L)(u) dS(K_1^{(m)}, \dots, K_{n-1}^{(m)}, u) \right. \\
&\quad \left. - \int_{S^{n-1}} (h_K - h_L)(u) dS(K_1, \dots, K_{n-1}, u) \right| \\
&\quad + \left| \int_{S^{n-1}} (f - (h_K - h_L))(u) dS(K_1, \dots, K_{n-1}, u) \right| \\
&\leq \|f - (h_K - h_L)\| n\, V(K_1 + B(1), \dots, K_{n-1} + B(1), B(1)) \\
&\quad + n |V(K_1^{(m)}, \dots, K_{n-1}^{(m)}, K) - V(K_1, \dots, K_{n-1}, K)| \\
&\quad + n |V(K_1^{(m)}, \dots, K_{n-1}^{(m)}, L) - V(K_1, \dots, K_{n-1}, L)| \\
&\quad + \|f - (h_K - h_L)\| n\, V(K_1, \dots, K_{n-1}, B(1)) \\
&\leq c(K_1, \dots, K_{n-1}) \varepsilon,
\end{aligned}$$

for $m \geq m_0$. □

Corollary 3.5.4. *For $j = 0, \dots, n-1$, the mapping $K \mapsto S_j(K, \cdot)$ on $\mathcal{K}^n$ is translation invariant, rotation covariant and continuous.*

Moreover,

$$S_{n-1}(K + B(\alpha), \cdot) = \sum_{j=0}^{n-1} \alpha^{n-1-j} \binom{n-1}{j} S_j(K, \cdot),$$

for $\alpha \geq 0$ (local STEINER *formula).*

Proof. We only have to prove the local STEINER formula. The latter follows from Theorem 3.5.3(a) and (b). □

The interpretation of the surface area measure $S_{n-1}(P, \cdot)$ for a polytope P is quite simple. For a Borel set $A \subset S^{n-1}$, the value of $S_{n-1}(P, A)$ gives the total surface area of the set of all boundary points of P which have an outer normal in A (since this set is a union of facets, the surface area is defined). In an appropriate way (and using approximation by polytopes), this interpretation carries over to arbitrary bodies K: $S_{n-1}(K, A)$ measures the total surface area of the set of all boundary points of K which have an outer normal in A. In particular, we have $S_{n-1}(K, \cdot) = 0$, if and only if $\dim K \leq n - 2$, and $S_{n-1}(K, \cdot) = V_{n-1}(K)(\varepsilon_u + \varepsilon_{-u})$, if $\dim K = n - 1$ and $K \perp u, u \in S^{n-1}$.

Now we study the problem, how far a convex body K is determined by one of its surface area measures $S_j(K, \cdot)$, $j \in \{1, \ldots, n - 1\}$. For $j = n - 1$ (and n-dimensional bodies), we can give a strong answer to this question.

Theorem 3.5.5. *Let* $K, L \in \mathcal{K}^n$ *with* $\dim K = \dim L = n$. *Then*

$$S_{n-1}(K, \cdot) = S_{n-1}(L, \cdot),$$

if and only if K *and* L *are translates.*

Proof. For translates K, L, the equality of the surface area measures follows from Corollary 3.5.4.

Assume now $S_{n-1}(K, \cdot) = S_{n-1}(L, \cdot)$. Then, Theorem 3.5.2 implies

$$\begin{aligned} V(K, \ldots, K, L) &= \frac{1}{n} \int_{S^{n-1}} h_L(u) dS_{n-1}(K, u) \\ &= \frac{1}{n} \int_{S^{n-1}} h_L(u) dS_{n-1}(L, u) \\ &= V(L). \end{aligned}$$

In the same way, we obtain $V(L, \ldots, L, K) = V(K)$. The MINKOWSKI inequalites (Theorem 3.4.3) therefore show that

$$V(L)^n \geq V(K)^{n-1} V(L)$$

and

$$V(K)^n \geq V(L)^{n-1} V(K),$$

which implies $V(K) = V(L)$. Therefore we have equality in both inequalities and hence K and L are homothetic. Since they have the same volume, they must be translates. □

The uniqueness result holds more generally for the j-th order surface area measures ($j \in \{1, \ldots, n - 1\}$), if the bodies have dimension at least $j + 1$ (for $j = 1$ even without a dimensional restriction). The proof uses a deep generalization of the MINKOWSKI inequalities (the ALEXANDROV-FENCHEL inequalities).

Theorem 3.5.5 can be used to express certain properties of convex bodies in terms of their surface area measures. We mention only one application of this type, other results can be found in the exercises. We recall that a convex body $K \in \mathcal{K}^n$ is *centrally symmetric*, if there is a point $x \in \mathbb{R}^n$ such that $K - x = -(K - x)$ (then $x \in K$ and x is the center of symmetry). Also, a measure μ on S^{n-1} is called *even*, if μ is invariant under reflection, i.e. $\mu(A) = \mu(-A)$, for all Borel sets $A \subset S^{n-1}$.

Corollary 3.5.6. *Let $K \in \mathcal{K}^n$ with $\dim K = n$. Then, K is centrally symmetric, if and only if $S_{n-1}(K, \cdot)$ is an even measure.*

In the following, we study the problem which measures μ on S^{n-1} arise as surface area measures $S_{n-1}(K, \cdot)$ of convex bodies K (the *existence problem*). Obviously, a necessary condition is that μ must have centroid 0. Another condition arises from a dimensional restriction. Namely, if $\dim K \leq n-2$, then $S_{n-1}(K, \cdot) = 0$, whereas for $\dim K = n-1, K \subset u^\perp, u \in S^{n-1}$, we have $S_{n-1}(K, \cdot) = V_{n-1}(K)(\varepsilon_u + \varepsilon_{-u})$ (both results follow from Theorem 3.5.2). Hence, for $\dim K \leq n-1$, the existence problem is not of any interest. Therefore, we now concentrate on bodies $K \in \mathcal{K}^n$ with $\dim K = n$. Again, Theorem 3.5.2 shows that this implies $\dim S_{n-1}(K, \cdot) = n$, where the latter condition means that $S_{n-1}(K, \cdot)$ is not supported by any lower dimensional sphere, i.e. $S_{n-1}(K, S^{n-1} \setminus E) > 0$ for all hyperplanes E through 0. As we shall show now, these two conditions (the centroid condition and the dimensional condition) characterize $(n-1)$-st surface area measures. We first prove the polytopal case.

Theorem 3.5.7. *For $k \geq n+1$, let $u_1, \dots, u_k \in S^{n-1}$ be unit vectors which span $\mathbb{R}^n$ and let $v^{(1)}, \dots, v^{(k)} > 0$ be numbers such that*

$$\sum_{i=1}^{k} v^{(i)} u_i = 0.$$

Then, there exists a (up to a translation unique) polytope $P \in \mathcal{P}^n$ with $\dim P = n$, for which

$$S_{n-1}(P, \cdot) = \sum_{i=1}^{k} v^{(i)} \varepsilon_{u_i},$$

i.e. the $u_1, \dots, u_k$ are the facet normals of P and the $v^{(1)}, \dots, v^{(k)}$ are the corresponding facet contents.

Proof. The uniqueness follows from Theorem 3.5.5.

For the existence, we denote by $\mathbb{R}^k_+$ the set of all vectors $y = (y^{(1)}, \dots, y^{(k)})$ with $y^{(i)} \geq 0, i = 1, \dots, k$. For $y \in \mathbb{R}^k_+$, let

$$P_{[y]} := \bigcap_{i=1}^{k} \{\langle \cdot, u_i \rangle \leq y^{(i)}\}.$$

Since $0 \in P_{[y]}$, this set is nonempty and polyhedral. Moreover, $P_{[y]}$ is bounded hence a convex polytope in $\mathbb{R}^n$. Namely, assuming $\alpha x \in P_{[y]}$, for some $x \in S^{n-1}$ and all $\alpha \geq 0$, we get

$$\langle x, u_i \rangle \leq 0, \qquad i = 1, \ldots, k.$$

Since the centroid condition implies

$$\sum_{i=1}^{k} v^{(i)} \langle x, u_i \rangle = 0$$

with $v^{(i)} > 0$ and $\langle x, u_i \rangle \leq 0$, it follows that

$$\langle x, u_1 \rangle = \cdots = \langle x, u_k \rangle = 0.$$

As a consequence $\langle x, z \rangle = 0$, for all $z \in \mathbb{R}^n$, since $u_1, \ldots, u_k$ span $\mathbb{R}^n$. Hence $x = 0$, a contradiction.

Therefore, $P_{[y]}$ is a polytope. We next show that the mapping $y \mapsto P_{[y]}$ is concave, i.e.

$$\gamma P_{[y]} + (1-\gamma) P_{[z]} \subset P_{[\gamma y + (1-\gamma) z]}, \tag{5.11}$$

for $y, z \in \mathbb{R}^k_+$ and $\gamma \in [0, 1]$. This follows since a point $x \in \gamma P_{[y]} + (1-\gamma) P_{[z]}$ satisfies $x = \gamma x' + (1-\gamma) x''$ with some $x' \in P_{[y]}$, $x'' \in P_{[z]}$, and hence

$$\langle x, u_i \rangle = \gamma \langle x', u_i \rangle + (1-\gamma) \langle x'', u_i \rangle \leq \gamma y^{(i)} + (1-\gamma) z^{(i)},$$

which shows that $x \in P_{[\gamma y + (1-\gamma) z]}$. Since the normal vectors u_i of the half spaces $\{\langle \cdot, u_i \rangle \leq y^{(i)}\}$ are fixed and only their distances $y^{(i)}$ from the origin vary, the mapping $y \mapsto P_{[y]}$ is continuous (with respect to the Hausdorff metric). Therefore, $y \mapsto V(P_{[y]})$ is continuous, which implies that the set

$$\mathcal{M} := \{y \in \mathbb{R}^k_+ : V(P_{[y]}) = 1\}$$

is nonempty and closed. The linear function

$$\varphi := \frac{1}{n} \langle \cdot, v \rangle, \qquad v := (v^{(1)}, \ldots, v^{(k)}),$$

is nonnegative on $\mathcal{M}$ (and continuous). Since $v^{(i)} > 0, i = 1, \ldots, k$, there is a vector y_0 such that $\varphi(y_0) =: \alpha \geq 0$ is the minimum of φ on $\mathcal{M}$. Since $y_0 \in \mathcal{M}$ implies $y_0^{(i)} > 0$ for some $i \in \{1, \ldots, k\}$, we get $\alpha > 0$.

We consider the polytope $Q := P_{[y_0]}$. Since $V(Q) = 1$, Q has interior points (and $0 \in Q$). We may assume that $0 \in \operatorname{int} Q$. Namely, for $0 \in \operatorname{bd} Q$, we can choose a translation vector $t \in \mathbb{R}^n$ such that $0 \in \operatorname{int}(Q + t)$. Then

$$Q + t = \bigcap_{i=1}^{k} \{\langle \cdot, u_i \rangle \leq \tilde{y}_0^{(i)}\}$$

with $\tilde{y}_0^{(i)} := y_0^{(i)} + \langle t, u_i \rangle$, $i = 1, \ldots, k$. Obviously, $\tilde{y}_0^{(i)} > 0$ and $Q + t = P_{[\tilde{y}_0]}$. Moreover, $V(Q+t) = V(Q) = 1$ and

$$\varphi(\tilde{y}_0) = \frac{1}{n}\langle y_0, v\rangle + \frac{1}{n}\sum_{i=1}^{k}\langle t, u_i\rangle v^{(i)} = \varphi(y_0) + \frac{1}{n}\langle t, \sum_{i=1}^{k} u_i v^{(i)}\rangle = \alpha,$$

since $\sum_{i=1}^{k} u_i v^{(i)} = 0$. Hence, we now assume $0 \in \operatorname{int} Q$, which gives us $y_0^{(i)} > 0$ for $i = 1, \ldots, k$. We define a vector $w = (w^{(1)}, \ldots, w^{(k)})$, where $w^{(i)} := V_{n-1}(Q(u_i))$ is the content of the support set of Q in direction u_i, $i = 1, \ldots, k$. Then,

$$\begin{aligned} 1 = V(Q) &= \frac{1}{n}\sum_{i=1}^{k} y_0^{(i)} w^{(i)} = \frac{1}{n}\langle y_0, w\rangle \\ &= \frac{1}{\alpha}\varphi(y_0) = \frac{1}{\alpha n}\langle y_0, v\rangle. \end{aligned}$$

Hence,

$$\langle y_0, w\rangle = \langle y_0, \frac{1}{\alpha} v\rangle = n.$$

Next, we define the hyperplanes

$$E := \{\langle \cdot, w\rangle = n\}$$

and

$$F := \{\langle \cdot, \frac{1}{\alpha} v\rangle = n\}$$

in $\mathbb{R}^k$. We want to show that $E = F$. First, we notice that $y_0 \in E \cap F$. Since y_0 has positive components, we can find a convex neighborhood U of y_0, such that $y \in U$ has the following two properties. First, $y^{(i)} > 0$ for $i = 1, \ldots, k$ and second every facet normal of $Q = P_{[y_0]}$ is also a facet normal of $P_{[y]}$. We now consider $y \in F \cap U$. Assume $V(P_{[y]}) > 1$, then there exists $0 < \beta < 1$ with

$$V(P_{[\beta y]}) = 1.$$

Since $y \in F$,

$$\varphi(\beta y) = \frac{1}{n}\langle \beta y, v\rangle = \beta\alpha < \alpha,$$

a contradiction. Therefore, $V(P_{[y]}) \leq 1$. For $\vartheta \in [0,1]$, the point $\vartheta y + (1-\vartheta) y_0$ is also in $F \cap U$. Therefore the volume inequality just proven applies and we get from (5.11)

$$V(\vartheta P_{[y]} + (1-\vartheta) Q) \leq V(P_{[\vartheta y + (1-\vartheta) y_0]}) \leq 1.$$

This yields

$$\begin{aligned} V(Q, \ldots, Q, P_{[y]}) &= \frac{1}{n}\lim_{\vartheta \to 0}\frac{V(\vartheta P_{[y]} + (1-\vartheta) Q) - (1-\vartheta)^n}{\vartheta} \\ &\leq \frac{1}{n}\lim_{\vartheta \to 0}\frac{1 - (1-\vartheta)^n}{\vartheta} \\ &= 1. \end{aligned}$$

Since by our assumption, each facet normal of Q is a facet normal of $P_{[y]}$, we have $h_{P_{[y]}}(u_i) = y^{(i)}$, for all i for which $w^{(i)} > 0$. Hence

$$1 \geq V(Q,\dots,Q,P_{[y]}) = \frac{1}{n}\sum_{i=1}^{k} h_{P_{[y]}}(u_i)w^{(i)} = \frac{1}{n}\langle y, w\rangle,$$

for all $y \in F \cap U$. This shows that $F \cap U \subset E$, which is only possible if $E = F$.

Since $E = F$ implies $w = \frac{1}{\alpha}v$, the polytope $P := \sqrt[n-1]{\alpha}Q$ fulfills all assertions of the theorem. □

We now extend this result to arbitrary bodies $K \in \mathcal{K}^n$.

Theorem 3.5.8. *Let μ be a finite Borel measure on S^{n-1} with centroid 0 and $\dim \mu = n$. Then, there exists a $($up to a translation unique$)$ body $K \in \mathcal{K}^n$, for which*

$$S_{n-1}(K,\cdot) = \mu.$$

Proof. Again, we only need to show the existence of K.

We make use of the fact that μ can be approximated (in the weak convergence) by discrete measures (measures with finite support) $\mu_j \to \mu$, for $j \to \infty$, which also have centroid 0 and fulfill $\dim \mu_j = n$. The measure μ_j can, for example, be constructed as follows. We divide S^{n-1} into finitely many pairwise disjoint Borel sets $A_{ij}, i = 0,1,\dots,k(j)$, such that $\mu(A_{0j}) = 0$, whereas $\operatorname{diam}(\operatorname{cl}\operatorname{conv} A_{ij}) < \frac{1}{j}$ and $\mu(A_{ij}) > 0$, for $i = 1,\dots,k(j)$. We then put

$$\mu_j := \sum_{i=1}^{k(j)} \mu(A_{ij})\|x_{ij}\|\varepsilon_{u_{ij}},$$

where

$$x_{ij} := \frac{1}{\mu(A_{ij})}\int_{A_{ij}} u\,d\mu(u),$$

and $u_{ij} := \frac{x_{ij}}{\|x_{ij}\|}$. This definition makes sense since, for $i \geq 1$, it can be shown that $0 \notin \operatorname{cl}\operatorname{conv} A_{ij}$ and therefore $x_{ij} \neq 0$. Moreover, μ_j has centroid 0 and converges to μ (see the exercises). Because of $\dim \mu = n$, we must have $\dim \mu_j = n$, for large enough j.

From Theorem 3.5.7, we obtain polytopes P_j with $0 \in P_j$ and

$$\mu_j = S_{n-1}(P_j,\cdot), \qquad j = 1,2,\dots.$$

We show that the sequence $(P_j)_{j\in\mathbb{N}}$ is uniformly bounded. First, $F(P_j) = \mu_j(S^{n-1}) \to \mu(S^{n-1})$ implies that

$$F(P_j) \leq C, \qquad j \in \mathbb{N},$$

for some $C > 0$. The isoperimetric inequality shows that then

$$V(P_j) \leq \tilde{C}, \qquad j \in \mathbb{N},$$

for another constant $\tilde{C} > 0$. Now let $x \in S^{n-1}$ and $\alpha_j \geq 0$ be such that $\alpha_j x \in P_j$, hence $[0, \alpha_j x] \subset P_j$. Since

$$h_{[0,\alpha_j x]} = \alpha_j \max(\langle x, \cdot \rangle, 0),$$

we get

$$\begin{aligned} V(P_j) &= \frac{1}{n} \sum_{i=1}^{k(j)} h_{P_j}(u_{ij}) V_{n-1}(P_j(u_{ij})) \\ &\geq \frac{1}{n} \sum_{i=1}^{k(j)} h_{[0,\alpha_j x]}(u_{ij}) V_{n-1}(P_j(u_{ij})) \\ &= \frac{\alpha_j}{n} \int_{S^{n-1}} \max(\langle x, u \rangle, 0) d\mu_j(u). \end{aligned}$$

The weak convergence implies

$$\frac{1}{n} \int_{S^{n-1}} \max(\langle x, u \rangle, 0) d\mu_j(u) \to \frac{1}{n} \int_{S^{n-1}} \max(\langle x, u \rangle, 0) d\mu(u),$$

and since both sides are support functions (as functions of x), the convergence is uniform in $x \in S^{n-1}$ (see Exercise 6 of Section 3.1). Because of $\dim \mu = n$ and since μ is centred, we get

$$f(x) := \frac{1}{n} \int_{S^{n-1}} \max(\langle x, u \rangle, 0) d\mu(u) > 0,$$

for all $x \in S^{n-1}$. As a support function, f is continuous, hence

$$c := \min_{x \in S^{n-1}} f(x)$$

exists and we have $c > 0$. Therefore, $\alpha_j \leq C'$ for all $j \geq j_0$, with a suitable $j_0 \in \mathbb{N}$ and a certain constant C'. This shows that the sequence $(P_j)_{j \in \mathbb{N}}$ is uniformly bounded.

By BLASCHKE's selection theorem, we can choose a convergent subsequence $P_{j_r} \to K$, $r \to \infty$, $K \in \mathcal{K}^n$. Then

$$S_{n-1}(P_{j_r}, \cdot) \to S_{n-1}(K, \cdot),$$

but also

$$S_{n-1}(P_{j_r}, \cdot) \to \mu.$$

Therefore, $S_{n-1}(K, \cdot) = \mu$. □

Exercises and problems

1. Let $K, M, L \in \mathcal{K}^n$ such that $K = M + L$. Show that

$$S_j(M,\cdot) = \sum_{i=0}^{j} (-1)^{j-i} \binom{j}{i} S(\underbrace{K,\dots,K}_{i}, \underbrace{L,\dots,L}_{j-i}, \underbrace{B(1),\dots,B(1)}_{n-1-j}, \cdot),$$

   for $j = 0, \dots, n-1$.

2. Let $K \in \mathcal{K}^n$ and $r(K)$ be the circumradius of K. Show that $r(K) \leq 1$ if and only if $V(K, M, \dots, M) \leq \frac{1}{n} F(M)$ for all $M \in \mathcal{K}^n$.

3. Let $\alpha \in (0,1)$ and $M, L \in \mathcal{K}^n$ with $\dim M = \dim L = n$.

   (a) Show that there is a convex body $K_\alpha \in \mathcal{K}^n$ with $\dim K_\alpha = n$ and

$$S_{n-1}(K_\alpha, \cdot) = \alpha S_{n-1}(M, \cdot) + (1-\alpha) S_{n-1}(L, \cdot).$$

   (b) Show that

$$V(K_\alpha)^{\frac{n-1}{n}} \geq \alpha V(M)^{\frac{n-1}{n}} + (1-\alpha) V(L)^{\frac{n-1}{n}},$$

   with equality if and only if M and L are homothetic.

4. Complete the proof of Theorem 3.5.8 by showing that the measures μ_j are well-defined (i.e. that $x_{ij} \neq 0$), have centroid 0, fulfill $\dim \mu_j = n$, for almost all j, and converge weakly to the given measure μ (as $j \to \infty$).

3.6 Projection functions

For a convex body $K \in \mathcal{K}^n$ and a direction $u \in S^{n-1}$, we define

$$v(K,u) := V_{n-1}(K \,|\, u^\perp),$$

the content of the orthogonal projection of K onto the hyperplane $u^\perp$. The function $v(K,\cdot)$ is called the *projection function* of K. We are interested in the information on the shape of K which can be deduced from the knowledge of its projection function $v(K,\cdot)$.

First, it is clear that translates K and $K+x, x \in \mathbb{R}^n$, have the same projection function. Second, K and $-K$ have the same projection function, which shows that in general K is not determined by $v(K,\cdot)$ (not even up to translations). The question occurs whether we get uniqueness up to translations and reflections. In order to give an answer, we need a representation of $v(K,\cdot)$.

Theorem 3.6.1. *For $K \in \mathcal{K}^n$ and $u \in S^{n-1}$, we have*

$$v(K,u) = \frac{1}{2}\int_{S^{n-1}} |\langle x,u\rangle| \, dS_{n-1}(K,x).$$

Proof. An application of FUBINI's theorem shows that

$$V(K+[-u,u]) = V(K) + 2v(K,u).$$

On the other hand, we have

$$V(K+[-u,u]) = \sum_{i=0}^{n}\binom{n}{i} V(\underbrace{K,\dots,K}_{i},\underbrace{[-u,u],\dots,[-u,u]}_{n-i}).$$

From Exercise 3.3.1, we know that

$$V(\underbrace{K,\dots,K}_{i},\underbrace{[-u,u],\dots,[-u,u]}_{n-i}) = 0,$$

for $i = 0,\dots,n-2$, hence

$$v(K,u) = \frac{n}{2}V(K,\dots,K,[-u,u]).$$

The assertion now follows from Theorem 3.5.2, since the segment $[-u,u]$ has support function $|\langle\cdot,u\rangle|$. $\square$

Remarks. A couple of properties of projection functions can be directly deduced from Theorem 3.6.1.

(1) We have $v(K,\cdot)=0$, if and only if $\dim K \le n-2$.

(2) If $\dim K = n-1$, $K \subset x^{\perp}$, then

$$v(K,\cdot) = V_{n-1}(K)|\langle x,\cdot\rangle|.$$

(3) If $\dim K = n$ and K is not centrally symmetric (i.e. $S_{n-1}(K,\cdot) \neq S_{n-1}(-K,\cdot)$), then there is an infinite family of bodies with the same projection function. Namely, for $\alpha \in [0,1]$, there is a body $K_\alpha \in \mathcal{K}^n$ with $\dim K_\alpha = n$ and

$$S_{n-1}(K_\alpha,\cdot) = \alpha S_{n-1}(K,\cdot) + (1-\alpha)S_{n-1}(-K,\cdot)$$

(this follows from Theorem 3.5.8). Then,

$$v(K_\alpha,\cdot) = \alpha v(K,\cdot) + (1-\alpha)v(-K,\cdot) = v(K,\cdot).$$

This also shows that there is always a centrally symmetric body, namely $K_{\frac{1}{2}}$, with the same projection function as K.

The body $K_{\frac{1}{2}}$ also has maximal volume in the class $\mathcal{C} := \{K_\alpha : \alpha \in [0,1]\}$ (by the BRUNN-MINKOWSKI theorem) and is moreover characterized by this fact; i.e., it is the only body in $\mathcal{C}$ with maximal volume.

(4) Since $|\langle x,\cdot\rangle|$ is a support function, the function $v(K,\cdot)$ is a positive combination of support functions, hence it is itself a support function of a convex body ΠK,

$$h_{\Pi K} := v(K,\cdot).$$

We call ΠK the *projection body* of K. The projection body is always centrally symmetric to the origin and, if $\dim K = n$, then $\dim \Pi K = n$.

Before we continue to discuss projection functions, we want to describe projection bodies geometrically.

Definition. A finite sum of segments $Z := s_1 + \cdots + s_k$ is called a *zonotope*. A *zonoid* is a convex body which is the limit (in the Hausdorff metric) of a sequence of zonotopes.

Zonotopes are polytopes and they are centrally symmetric. Namely, if $s_i = [-y_i, y_i] + x_i$ is the representation of the segment s_i (with center x_i and endpoints $-y_i + x_i$, $y_i + x_i$), then

$$Z = \sum_{i=1}^{k}[-y_i, y_i] + \sum_{i=1}^{k} x_i.$$

Hence, $x := \sum_{i=1}^{k} x_i$ is the center of Z. Zonoids, as limits of zonotopes, are also centrally symmetric. We assume w.l.o.g. that the center of zonotopes and zonoids is the origin and denote the correspondings set of zonoids by $\mathcal{Z}^n$.

The following results show that zonoids and projection bodies are closely related.

Theorem 3.6.2. *Let $K \in \mathcal{K}^n$. Then, K is a zonoid, if and only if there exists an even Borel measure $\mu(K,\cdot)$ on S^{n-1} such that*

$$h_K(u) = \int_{S^{n-1}} |\langle x, u\rangle| d\mu(K,x).$$

For a zonoid K, such a measure $\mu(K,\cdot)$ is called a *generating measure* of K. We shall soon see that $\mu(K,\cdot)$ is uniquely determined by h_K.

Proof. Suppose

$$h_K(u) = \int_{S^{n-1}} |\langle x, u\rangle| d\mu(K,x).$$

As in the proof of Theorem 3.5.8, we find a sequence of even, discrete measures $\mu_j \to \mu(K,\cdot)$,

$$\mu_j := \frac{1}{2}\sum_{i=1}^{k(j)} \alpha_{ij}(\epsilon_{u_{ij}} + \epsilon_{-u_{ij}}), \qquad u_{ij} \in S^{n-1}, \alpha_{ij} > 0.$$

Then,

$$Z_j := \sum_{i=1}^{k(j)} [-\alpha_{ij}u_{ij}, \alpha_{ij}u_{ij}]$$

is a zonotope and

$$\begin{aligned} h_{Z_j}(u) &= \int_{S^{n-1}} |\langle x, u\rangle| d\mu_j(x) \\ &\to \int_{S^{n-1}} |\langle x, u\rangle| d\mu(K,x) = h_K(u), \end{aligned}$$

for all $u \in S^{n-1}$. Therefore, $Z_j \to K$ (as $j \to \infty$), i.e. K is a zonoid.

Conversely, assume that $K = \lim_{j\to\infty} Z_j$, Z_j zonotope. Then,

$$Z_j = \sum_{i=1}^{k(j)} [-y_{ij}, y_{ij}]$$

with suitable vectors $y_{ij} \in \mathbb{R}^n$. Consequently,

$$\begin{aligned} h_{Z_j}(u) &= \sum_{i=1}^{k(j)} |\langle y_{ij}, u\rangle| \\ &= \int_{S^{n-1}} |\langle x, u\rangle| d\mu_j(x), \end{aligned}$$

where

$$\mu_j := \frac{1}{2}\sum_{i=1}^{k(j)} \|y_{ij}\|(\epsilon_{u_{ij}} + \epsilon_{-u_{ij}})$$

and

$$u_{ij} := \frac{y_{ij}}{\|y_{ij}\|}.$$

We would like to show that the sequence $(\mu_j)_{j\in\mathbb{N}}$ converges weakly.

We have

$$\int_{S^{n-1}} h_{Z_j}(u)du = \kappa_{n-1}V_1(Z_j) \to \kappa_{n-1}V_1(K).$$

Also, using FUBINI's theorem and Theorem 3.6.1 (for the unit ball), we get

$$\int_{S^{n-1}} h_{Z_j}(u)du = \int_{S^{n-1}}\int_{S^{n-1}} |\langle x,u\rangle| du d\mu_j(x) = 2\kappa_{n-1}\mu_j(S^{n-1}).$$

Hence, $\mu_j(S^{n-1})$ is bounded from above by a constant C, for all j. Now we use the fact that the set $\mathcal{M}_C$ of all Borel measures ρ on S^{n-1} with $\rho(S^{n-1}) \le C$ is weakly compact (see, e.g., the books of Billingsley, *Convergence of probability measures*, Wiley 1968, p. 37; or Gänssler-Stute, *Wahrscheinlichkeitstheorie*, Springer 1977, p. 344). Therefore, $(\mu_j)_{j\in\mathbb{N}}$ contains a convergent subsequence. W.l.o.g., we may assume that $(\mu_j)_{j\in\mathbb{N}}$ converges to a limit measure which we denote by $\mu(K,\cdot)$. The weak convergence implies that

$$\begin{aligned} h_K(u) &= \lim_{j\to\infty} h_{Z_j}(u) \\ &= \lim_{j\to\infty}\int_{S^{n-1}} |\langle x,u\rangle| d\mu_j(x) \\ &= \int_{S^{n-1}} |\langle x,u\rangle| d\mu(K,x). \end{aligned}$$

□

Remark. As the above proof shows, we have $\dim K = n$, if and only if $\dim \mu(K,\cdot) = n$.

Corollary 3.6.3. *The projection body* ΠK *of a convex body* K *is a zonoid. Reversely, if* Z *is a zonoid with* $\dim Z = n$*, then there is a convex body* K *with* $\dim K = n$ *and which is centrally symmetric to the origin and fulfills*

$$Z = \Pi K.$$

Proof. The first result follows from Theorems 3.6.1 and 3.6.2. For the second, Theorem 3.6.2 shows that

$$h_Z(u) = \int_{S^{n-1}} |\langle x,u\rangle| d\mu(Z,x)$$

with an even measure $\mu(Z,\cdot)$, $\dim \mu(Z,\cdot) = n$. By Theorem 3.5.8,

$$\mu(Z,\cdot) = S_{n-1}(K,\cdot),$$

for some convex body $K \in \mathcal{K}^n$, $\dim K = n$, and hence $Z = \Pi K$. By Corollary 3.5.6, K is centrally symmetric. $\square$

Finally, we want to show that the generating measure of a zonoid is uniquely determined. We first need two auxiliary lemmas. If A is the $(n \times n)$-matrix of an injective linear mapping in $\mathbb{R}^n$, we define

$$AZ := \{Ax : x \in Z\}$$

and denote by $A\mu$, for a measure μ on S^{n-1}, the image measure of

$$\int_{(\cdot)} \|Ax\| d\mu(x)$$

under the mapping

$$x \mapsto \frac{Ax}{\|Ax\|}, \quad x \in S^{n-1}.$$

Lemma 3.6.4. *If $Z \in \mathcal{K}^n$ is a zonoid and*

$$h_Z = \int_{S^{n-1}} |\langle x, \cdot \rangle| d\mu(Z, x),$$

then AZ is a zonoid and

$$h_{AZ} = \int_{S^{n-1}} |\langle x, \cdot \rangle| dA\mu(Z, x).$$

Proof. We have

$$\begin{aligned} h_{AZ} &= \sup_{x \in AZ} \langle u, x \rangle = \sup_{x \in Z} \langle u, Ax \rangle = \sup_{x \in Z} \langle A^T u, x \rangle = h_Z(A^T u) \\ &= \int_{S^{n-1}} |\langle x, A^T u \rangle| d\mu(Z, x) = \int_{S^{n-1}} |\langle Ax, u \rangle| d\mu(Z, x) \\ &= \int_{S^{n-1}} |\langle \frac{Ax}{\|Ax\|}, u \rangle| \|Ax\| d\mu(Z, x) = \int_{S^{n-1}} |\langle y, u \rangle| dA\mu(Z, y). \end{aligned}$$

$\square$

Let $\mathcal{V}$ denote the vector space of functions

$$f = \int_{S^{n-1}} |\langle x, \cdot \rangle| d\mu(x) - \int_{S^{n-1}} |\langle x, \cdot \rangle| d\rho(x),$$

where μ, ρ vary among all finite even Borel measures on S^{n-1}. $\mathcal{V}$ is a subspace of the Banach space $\mathbf{C}_e(S^{n-1})$ of even continuous functions on S^{n-1}.

Lemma 3.6.5. *The vector space $\mathcal{V}$ is dense in $\mathbf{C}_e(S^{n-1})$.*

Proof. Choosing $\mu = c\omega_{n-1}$, for $c \geq 0$, and $\rho = 0$ (or vice versa), we see that $\mathcal{V}$ contains all constant functions.

By Lemma 3.6.4, the support functions $h_{AB(1)}$ lie in $\mathcal{V}$, for all regular $(n \times n)$-matrices A (the body $AB(1)$ is an ellipsoid). Since

$$h_{AB(1)}(u) = \|A^T u\|, \qquad u \in S^{n-1},$$

we obtain all functions

$$\begin{aligned} f(B,u) &:= \sqrt{\langle Au, Au\rangle} = \sqrt{\langle A^T A u, u\rangle} \\ &= \sqrt{\langle Bu, u\rangle} = \left(\sum_{i,j=1}^{n} b_{ij} u_i u_j\right)^{\frac{1}{2}}, \end{aligned}$$

where $B = A^T A = ((b_{ij}))$ varies among the positive definite symmetric $(n \times n)$-matrices B. Here, in deviation of our usual notation, we used $u = (u_1, ..., u_n)$ and we also consider $f(B,u)$, for fixed u and in view of the symmetry of B, as a function of the $n(n+1)/2$ variables $b_{ij}, 1 \leq i \leq j \leq n$. For $\epsilon > 0$ and $1 \leq i_0 \leq j_0 \leq n$, let $\tilde{B} = ((\tilde{b}_{ij}))$ with

$$\tilde{b}_{ij} := \begin{cases} b_{ij} + \epsilon & \text{if} \quad (i,j) \in \{(i_0, j_0), (j_0, i_0)\}, \\ b_{ij} & \phantom{\text{if}} \quad (i,j) \notin \{(i_0, j_0), (j_0, i_0)\}. \end{cases}$$

Then, $\tilde{B}$ is symmetric and positive definite, for small enough ϵ. Consequently,

$$\frac{f(\tilde{B},\cdot) - f(B,\cdot)}{\epsilon} \in \mathcal{V}$$

and

$$\lim_{\epsilon \to 0} \frac{f(\tilde{B},\cdot) - f(B,\cdot)}{\epsilon} = \frac{\partial f}{\partial b_{i_0, j_0}}(B,\cdot) \in \operatorname{cl} \mathcal{V}.$$

A direct computation yields

$$\frac{\partial f}{\partial b_{i_0 j_0}}(B,u) = \frac{u_{i_0} u_{j_0}}{f(B,u)}, \quad \text{for } i_0 < j_0,$$

(and $\frac{\partial f}{\partial b_{i_0 i_0}}(B,u) = \frac{u_{i_0}^2}{2f(B,u)}$). Repeating this argument with $b_{i_1 j_1}$ etc., we obtain that all partial derivatives of f w.r.t. the variables $b_{ij}, 1 \leq i \leq j \leq n$, are in $\operatorname{cl} \mathcal{V}$, hence all functions

$$u \mapsto \frac{u_1^{i_1} \cdots u_n^{i_n}}{f(B,u)^k}, \qquad i_1 + \cdots + i_n = 2k, k = 1, 2, \ldots.$$

Now we choose B to be the unit matrix. Then $f(B,\cdot) = 1$, hence all even polynomials are in $\operatorname{cl} \mathcal{V}$. The theorem of STONE-WEIERSTRASS now shows that $\operatorname{cl} \mathcal{V} = \mathbf{C}_e(S^{n-1})$. □

Theorem 3.6.6. *For a zonoid $Z \in \mathcal{K}^n$, the generating measure is uniquely determined.*

Proof. Assume we have two even measures $\mu := \mu(Z, \cdot)$ and ρ on S^{n-1} with

$$\int_{S^{n-1}} |\langle x, \cdot\rangle| d\mu(x) = \int_{S^{n-1}} |\langle x, \cdot\rangle| d\rho(x).$$

Then,

$$\int_{S^{n-1}} \int_{S^{n-1}} |\langle x, u\rangle| d\mu(x) d\tilde{\mu}(u) = \int_{S^{n-1}} \int_{S^{n-1}} |\langle x, u\rangle| d\rho(x) d\tilde{\mu}(u),$$

for all measures $\tilde{\mu}$ on S^{n-1}. Replacing $\tilde{\mu}$ by a difference of measures and applying FUBINI's theorem, we obtain

$$\int_{S^{n-1}} f(x) d\mu(x) = \int_{S^{n-1}} f(x) d\rho(x),$$

for all functions $f \in \mathcal{V}$. Lemma 3.6.5 shows that this implies $\mu = \rho$. □

Combining Theorem 3.6.6 with Theorems 3.6.1, 3.6.2 and 3.5.5, we get directly our final result in this chapter.

Corollary 3.6.7. *A centrally symmetric convex body $K \in \mathcal{K}^n$ with $\dim K = n$ is uniquely determined (up to translations) by its projection function $v(K, \cdot)$.*

Exercises and problems

1. Let $n \geq 3$ and $P \in \mathcal{K}^n$ be a polytope. Show that P is a zonotope, if and only if all 2-faces of P are centrally symmetric.

2. Let $Z \in \mathcal{K}^n$ be a zonoid and $u_1, \dots, u_k \in S^{n-1}$.

   (a) Show that there exists a zonotope P such that

   $$h_Z(u_i) = h_P(u_i), \qquad i = 1, \dots, k.$$

   Hint: Use CARATHÉODORY's theorem for a suitable subset A of $\mathbb{R}^k$.

   (b) Show in addition that P can be chosen to be the sum of at most k segments.

   Hint: Replace CARATHÉODORY's theorem by the theorem of BUNDT.

3. Let $P, Q \in \mathcal{P}^n$ be zonotopes and $K \in \mathcal{K}^n$ a convex body such that

   $$P = K + Q.$$

   Show that K is also a zonotope.

∗ 4. Let $P \in \mathcal{P}^n$ be a polytope. Show that P is a zonotope, if and only if h_P fulfills the HLAWKA inequality:

$$(*) \qquad h_P(x) + h_P(y) + h_P(z) + h_P(x+y+z) \geq h_P(x+y) + h_P(x+z) + h_P(y+z),$$

for all $x, y, z \in \mathbb{R}^n$.

Hint: For one direction, show first that $(*)$ implies the central symmetry of P and then that $(*)$ implies the HLAWKA inequality for each face $P(u), u \in S^{n-1}$. Then use Exercise 1 above.

5. Let $Z \in \mathcal{K}^n$ be a zonoid.

   (a) For $u \in S^{n-1}$, show that $Z(u)$ is a zonoid and that

   $$h_{Z(u)} = \int_{S^{n-1} \cap u^\perp} |\langle x, \cdot \rangle| \mu(Z, dx) + \langle x_u, \cdot \rangle,$$

   where

   $$x_u := 2 \int_{\{x \in S^{n-1} : \langle x, u \rangle > 0\}} x d\mu(Z, x).$$

   (b) Use (a) to show that a zonoid which is a polytope must be a zonotope.

Chapter 4

Integral geometric formulas

In this final chapter, we discuss integral formulas for intrinsic volumes $V_j(K)$, which are based on sections and projections of convex bodies K. We shall also discuss some applications of stereological nature.

As a motivation, we start with the formula for the projection function $v(K, \cdot)$ from Theorem 3.6.1. Integrating $v(K, u)$ over all $u \in S^{n-1}$ (with respect to the spherical LEBESGUE measure ω_{n-1}), we obtain

$$\int_{S^{n-1}} v(K, u) du = \kappa_{n-1} \int_{S^{n-1}} dS_{n-1}(K, x) = 2\kappa_{n-1} V_{n-1}(K).$$

Since $v(K, u) = V_{n-1}(K|u^\perp)$, we may replace the integration over S^{n-1} by one over the space $\mathcal{L}_{n-1}^n$ of hyperplanes (through 0) in $\mathbb{R}^n$, namely by considering the normalized image measure ν_{n-1} of ω_{n-1} under the mapping $u \mapsto u^\perp$. Denoting the integration by ν_{n-1} shortly as dL_{n-1}, we then get

$$\int_{\mathcal{L}_{n-1}^n} V_{n-1}(K|L_{n-1}) dL_{n-1} = \frac{2\kappa_{n-1}}{n\kappa_n} V_{n-1}(K).$$

This is known as CAUCHY's surface area formula for convex bodies. Our first goal is to generalize this projection formula to other intrinsic volumes V_j and to projection flats L_q of lower dimensions. This requires a natural measure ν_q on the space of q-dimensional subspaces first. Later we will also consider integrals over sections of K with affine flats and integrate those with a natural measure μ_q on (affine) q-flats. The first section discusses how the measures ν_q and μ_q can be introduced in an elementary way.

4.1 Invariant measures

We begin with the set $\mathcal{L}_q^n$ of q-dimensional (linear) subspaces of $\mathbb{R}^n$, $q \in \{0, \dots, n-1\}$. $\mathcal{L}_q^n$ becomes a compact metric space, if we define the distance $d(L, L')$, for $L, L' \in \mathcal{L}_q^n$, as the HAUSDORFF distance of $L \cap B(1)$ and $L' \cap B(1)$. We want to introduce an *invariant probability measure* ν_q on $\mathcal{L}_q^n$. Here, probability measure refers to the Borel σ-algebra generated by the metric structure and invariance refers to the rotation group SO_n and means that

$$\nu_q(\vartheta\mathcal{A}) = \nu_q(\mathcal{A}),$$

for all $\vartheta \in SO_n$ and all Borel sets $\mathcal{A} \subset \mathcal{L}_q^n$ (with $\vartheta\mathcal{A} := \{\vartheta L : L \in \mathcal{A}\}$). We will obtain ν_q as the image measure of an invariant measure ν on SO_n.

The rotation group SO_n can be viewed as a subset of $(S^{n-1})^n \subset \mathbb{R}^{n^2}$, if we identify rotations ϑ with orthogonal matrices A (with $\det A = 1$) and then replace A by the n-tuple $(a_1, \dots, a_n) \in (S^{n-1})^n$ of column vectors. The euclidean metric on $\mathbb{R}^{n^2}$ therefore induces a metric on SO_n and SO_n becomes a compact metric space in this way. It is easy to see that the operations of multiplication and inversion in SO_n (i.e. the mappings $(\vartheta, \eta) \mapsto \vartheta\eta$ and $\vartheta \mapsto \vartheta^{-1}$) are continuous. This shows that SO_n (with the given metric) is a compact *topological group*. A general theorem in the theory of topological groups implies the existence and uniqueness of an invariant probability measure ν on SO_n (called the HAAR measure). Since SO_n is not commutative, invariance means here

$$\nu(\vartheta\mathcal{A}) = \nu(\mathcal{A}), \qquad \nu(\mathcal{A}\vartheta) = \nu(\mathcal{A}), \qquad \nu(\mathcal{A}^{-1}) = \nu(\mathcal{A}),$$

for all $\vartheta \in SO_n$ and all Borel sets $\mathcal{A} \subset SO_n$, where

$$\vartheta\mathcal{A} := \{\vartheta\eta : \eta \in \mathcal{A}\}, \quad \mathcal{A}\vartheta := \{\eta\vartheta : \eta \in \mathcal{A}\}, \quad \mathcal{A}^{-1} := \{\eta^{-1} : \eta \in \mathcal{A}\}.$$

However, we can show the existence of ν also by a direct construction.

Lemma 4.1.1. *There is an invariant probability measure ν on SO_n.*

Proof. We consider the set $LU_n \subset (S^{n-1})^n$ of linearly independent n-tuples. LU_n is open and the complement has measure zero with respect to $\omega_{n-1} \otimes \cdots \otimes \omega_{n-1}$. On LU_n we define the mapping T onto SO_n by

$$T(x_1, \dots, x_n) := \left(\frac{y_1}{\|y_1\|} \vdots \cdots \vdots \frac{y_n}{\|y_n\|} \right), \tag{1.1}$$

where $(y_1, \dots, y_n)$ is the n-tuple obtained from $(x_1, \dots, x_n)$ by the GRAM-SCHMIDT orthogonalization procedure (and where, in addition, the sign of y_n is chosen such that the matrix on the right side of (1.1) has determinant 1). Up to the sign of y_n, we thus have

$$y_k := x_k - \sum_{i=1}^{k-1} \langle x_k, y_i \rangle \frac{y_i}{\|y_i\|^2}, \qquad k = 2, \dots, n,$$

and $y_1 := x_1$. T is almost everywhere defined (with respect to $\omega_{n-1} \otimes \cdots \otimes \omega_{n-1}$) and continuous. Let $\overline{\nu}$ be the image measure of $\omega_{n-1} \otimes \cdots \otimes \omega_{n-1}$ under T. For each continuous function f on SO_n and $\vartheta \in SO_n$, we then get

$$\begin{aligned}
\int_{SO_n} f(\vartheta\eta) d\overline{\nu}(\eta) &= \int_{S^{n-1}} \cdots \int_{S^{n-1}} f(\vartheta T(x_1, \dots, x_n)) dx_1 \cdots dx_n \\
&= \int_{S^{n-1}} \cdots \int_{S^{n-1}} f(\left(\frac{\vartheta y_1}{\|y_1\|} \vdots \cdots \vdots \frac{\vartheta y_n}{\|y_n\|} \right)) dx_1 \cdots dx_n.
\end{aligned}$$

Obviously,

$$\left(\frac{\vartheta y_1}{\|y_1\|}\vdots\cdots\vdots\frac{\vartheta y_n}{\|y_n\|}\right) = T(\vartheta x_1, \dots, \vartheta x_n),$$

and we obtain

$$\int_{SO_n} f(\vartheta\eta)d\overline{\nu}(\eta) = \int_{SO_n} f(\eta)d\overline{\nu}(\eta).$$

This shows that $\overline{\nu}$ is invariant from the left.

For the inversion invariance, we first observe

$$\begin{aligned}\int_{SO_n} f(\eta^{-1}\vartheta)d\overline{\nu}(\eta) &= \int_{SO_n} f((\vartheta^{-1}\eta)^{-1})d\overline{\nu}(\eta) = \int_{SO_n} g(\vartheta^{-1}\eta)d\overline{\nu}(\eta)\\ &= \int_{SO_n} g(\eta)d\overline{\nu}(\eta) = \int_{SO_n} f(\eta^{-1})d\overline{\nu}(\eta),\end{aligned} \tag{1.2}$$

for continuous f, g and all $\vartheta \in SO_n$, where $g(\rho) = f(\rho^{-1}), \rho \in SO_n$, and where we used the left invariance of $\overline{\nu}$. Hence, by FUBINI's theorem,

$$\int_{SO_n} f(\eta^{-1})d\overline{\nu}(\eta) = \int_{SO_n}\int_{SO_n} f(\eta^{-1}\vartheta)d\overline{\nu}(\vartheta)d\overline{\nu}(\eta) = \int_{SO_n} f(\vartheta)d\overline{\nu}(\vartheta),$$

again from the left invariance.

Finally, the right invariance follows from (1.2) and the inversion invariance,

$$\int_{SO_n} f(\eta\vartheta)d\overline{\nu}(\eta) = \int_{SO_n} f(\eta^{-1}\vartheta)d\overline{\nu}(\eta) = \int_{SO_n} f(\eta^{-1})d\overline{\nu}(\eta) = \int_{SO_n} f(\eta)d\overline{\nu}(\eta).$$

The normalized measure $\nu := (1/n\kappa_n)^n\overline{\nu}$ fulfills now all assertions of the lemma. □

For the rest of this chapter we choose a fixed subspace $L_q^0 \in \mathcal{L}_q^n$ as a reference space and define

$$\nu_q := \Phi \circ \nu,$$

where

$$\Phi : SO_n \to \mathcal{L}_q^n, \qquad \vartheta \mapsto \vartheta L_q^0.$$

It is easy to see that Φ is continuous and therefore measurable. The definition of ν_q is based on the fact that the rotation group SO_n operates *transitively* on $\mathcal{L}_q^n$, which means that for given $L, L' \in \mathcal{L}_q^n$ there is always a rotation ϑ with $L' = \vartheta L$. This implies that the images $\vartheta L_q^0, \vartheta \in SO_n$, run through all elements of $\mathcal{L}_q^n$. We abbreviate the integration with respect to ν_q by dL_q. Then,

$$\int_{\mathcal{L}_q^n} f(L_q)dL_q = \int_{SO_n} f(\vartheta L_q^0)d\nu(\vartheta),$$

for all continuous functions f on $\mathcal{L}_q^n$.

Theorem 4.1.2. *For $q \in \{1, \dots, n-1\}$, the measure ν_q is an invariant probability measure. It is the only invariant probability measure on $\mathcal{L}_q^n$.*

Moreover, for a continuous function f on $\mathcal{L}_q^n$, we have

$$\int_{\mathcal{L}_q^n} f(L_q)dL_q = \int_{\mathcal{L}_{n-q}^n} f(L_{n-q}^\perp)dL_{n-q},$$

for $1 \le q \le n-1$, and

$$\int_{\mathcal{L}_q^n} f(L_q)dL_q = \int_{\mathcal{L}_m^n} \left(\int_{\mathcal{L}_q^n(L_m)} f(L_q)dL_q \right) dL_m,$$

for $0 \le q < m \le n-1$. (Here $\mathcal{L}_q^n(L_m) := \{L_q \in \mathcal{L}_q^n : L_q \subset L_m\}$ and we identify this set with $\mathcal{L}_q^m$.)

Proof. Obviously, ν_q is a probability measure. To show its invariance, let f be a continuous function on $\mathcal{L}_q^n$ and $\eta \in SO_n$. Then

$$\begin{aligned}\int_{\mathcal{L}_q^n} f(\eta L_q)dL_q &= \int_{SO_n} f(\eta\vartheta L_q^0)d\nu(\vartheta)\\ &= \int_{SO_n} f(\rho L_q^0)d\nu(\rho)\\ &= \int_{\mathcal{L}_q^n} f(L_q)dL_q.\end{aligned}$$

Next, we show the uniqueness. Assume that ν_q' is also an invariant probability measure on $\mathcal{L}_q^n$. Then

$$\begin{aligned}\int_{\mathcal{L}_q^n} f(L_q)d\nu_q'(L_q) &= \int_{SO_n}\int_{\mathcal{L}_q^n} f(\vartheta L_q)d\nu_q'(L_q)d\nu(\vartheta)\\ &= \int_{\mathcal{L}_q^n}\int_{SO_n} f(\vartheta L_q)d\nu(\vartheta)d\nu_q'(L_q)\end{aligned}$$

For $L_q \in \mathcal{L}_q^n$ there exists an $\eta \in SO_n$ with $L_q = \eta L_q^0$, hence

$$\begin{aligned}\int_{SO_n} f(\vartheta L_q)d\nu(\vartheta) &= \int_{SO_n} f(\vartheta\eta L_q^0)d\nu(\vartheta)\\ &= \int_{SO_n} f(\vartheta L_q^0)d\nu(\vartheta).\end{aligned}$$

This shows that the function

$$L_q \mapsto \int_{SO_n} f(\vartheta L_q)d\nu(\vartheta)$$

is a constant $c(f)$, which implies that

$$\int_{\mathcal{L}_q^n} f(L_q)d\nu_q'(L_q) = c(f)\int_{\mathcal{L}_q^n} d\nu_q'(L_q) = c(f).$$

In the same way, we get

$$\int_{\mathcal{L}_q^n} f(L_q)dL_q = c(f),$$

hence

$$\int_{\mathcal{L}_q^n} f(L_q)d\nu_q'(L_q) = \int_{\mathcal{L}_q^n} f(L_q)dL_q,$$

for all continuous functions f on $\mathcal{L}_q^n$. Therefore, $\nu_q' = \nu_q$.

The two integral formulas now follow from the uniqueness of ν_q. Namely,

$$f \mapsto \int_{\mathcal{L}_{n-q}^n} f(L_{n-q}^\perp)dL_{n-q}$$

defines a probability measure ν_q' on $\mathcal{L}_q^n$ (by the RIESZ representation theorem). The invariance of ν_{n-q} shows that ν_q' is invariant, hence $\nu_q' = \nu_q$. In the same manner, we obtain the second, iterated integral formula. Here, the uniqueness result is already used to show that the invariant measure $\nu_q^{\vartheta L_m}$ on $\mathcal{L}_q^n(\vartheta L_m)$ is the image under $L \mapsto \vartheta L$ of the invariant measure $\nu_q^{L_m}$ on $\mathcal{L}_q^n(L_m)$. □

Now we consider the set $\mathcal{E}_q^n$ of affine q-dimensional subspaces (q-*flats*, for short) in $\mathbb{R}^n$. Each $E_q \in \mathcal{E}_q^n$ has a unique representation $E_q = L_q + x$, $L_q \in \mathcal{L}_q^n, x \in L_q^\perp$. This allows us to define a metric on $\mathcal{E}_q^n$, namely as

$$d(E_q, E_q') := d(L_q, L_q') + d(x, x').$$

The metric space $\mathcal{E}_q^n$ is locally compact but not compact. We define the measure μ_q as the image

$$\mu_q := \Psi \circ (\nu \otimes \lambda_{n-q}),$$

where

$$\Psi : SO_n \times (L_q^0)^\perp \to \mathcal{E}_q^n, \qquad (\vartheta, x) \mapsto \vartheta(L_q^0 + x)$$

and λ_{n-q} is the LEBESGUE measure on $(L_q^0)^\perp$. Apparently, $\mu_q(\mathcal{E}_q^n) = \infty$, but the set $\mathcal{E}_q^n(B(1))$ of q-flats intersecting the unit ball has finite measure,

$$\mu_q(\mathcal{E}_q^n(B(1))) = \kappa_{n-q}.$$

For the measure μ_q, invariance refers to the group G_n of rigid motions, that is

$$\mu_q(g\mathcal{A}) = \mu_q(\mathcal{A}),$$

for all $g \in G_n$ and all Borel sets $\mathcal{A} \subset \mathcal{E}_q^n$ (again $g\mathcal{A} := \{gL : L \in \mathcal{A}\}$). As in the case of ν_q, we will denote integration by μ_q simply as dE_q. For a flat $E_m \in \mathcal{E}_m^n, q < m \leq n-1$, we put $\mathcal{E}_q^n(E_m) := \{E_q \in \mathcal{E}_q^n : E_q \subset E_m\}$. Because of the unique decomposition $E_m = L_m + x$, $L_m \in \mathcal{L}_m^n, x \in L_m^\perp$, we may identify $\mathcal{E}_q^n(E_m)$ with $\mathcal{E}_q^m$ (by mapping x to the origin). We denote by dL_m the integration with respect to the corresponding measure on $\mathcal{E}_q^n(E_m)$.

Theorem 4.1.3. *For $q \in \{0, \dots, n-1\}$, μ_q is an invariant measure.*
For a continuous function f on $\mathcal{E}_q^n$ with compact support, we have

$$\int_{\mathcal{E}_q^n} f(E_q) dE_q = \int_{\mathcal{L}_q^n} \int_{L_q^\perp} f(L_q + x) dx dL_q.$$

Furthermore,

$$\int_{\mathcal{E}_q^n} f(E_q) dE_q = \int_{\mathcal{E}_m^n} \left(\int_{\mathcal{E}_q^n(E_m)} f(E_q) dE_q \right) dE_m,$$

for $0 \le q < m \le n-1$.

Proof. For the invariance of μ_q, we consider $g \in G_n$ and a continuous function f on $\mathcal{E}_q^n$ with compact support. By definition of μ_q,

$$\int_{\mathcal{E}_q^n} f(gE_q) dE_q = \int_{SO_n} \int_{(L_q^0)^\perp} f(g\vartheta(L_q^0 + x)) dx d\nu(\vartheta).$$

We decompose g into rotation and translation,

$$g : z \mapsto \eta(z + y), \qquad \eta \in SO_n, y \in \mathbb{R}^n,$$

and put $x' := \vartheta^{-1} y | (L_q^0)^\perp$. Then,

$$g\vartheta(L_q^0 + x) = \eta\vartheta(L_q^0 + x + x'),$$

hence

$$\begin{aligned} \int_{\mathcal{E}_q^n} f(gE_q) dE_q &= \int_{SO_n} \int_{(L_q^0)^\perp} f(\rho(L_q^0 + z)) dz d\nu(\rho) \\ &= \int_{\mathcal{E}_q^n} f(E_q) dE_q. \end{aligned}$$

The first integral formula follows from

$$\begin{aligned} \int_{\mathcal{E}_q^n} f(E_q) dE_q &= \int_{SO_n} \int_{(L_q^0)^\perp} f(\vartheta(L_q^0 + x)) dx d\nu(\vartheta) \\ &= \int_{SO_n} \int_{(\vartheta L_q^0)^\perp} f(\vartheta L_q^0 + x)) dx d\nu(\vartheta) \\ &= \int_{\mathcal{L}_q^n} \int_{L_q^\perp} f(L_q + x) dx dL_q. \end{aligned}$$

For the second integral formula, we consider $E_q \in \mathcal{E}_q^n(E_m)$. Because of $E_m = L_m + x$, $L_m \in \mathcal{L}_m^n$, $x \in L_m^\perp$, we get

$$E_q = L_q + x + y$$

with $L_q \in \mathcal{L}_q^n(L_m)$ and $y \in L_q^\perp \cap L_m$. Therefore,

$$\int_{\mathcal{E}_m^n}\left(\int_{\mathcal{E}_q^n(E_m)} f(E_q)dE_q\right)dE_m$$

$$= \int_{\mathcal{L}_m^n}\int_{L_m^\perp}\left(\int_{\mathcal{L}_q^n(L_m)}\int_{L_q^\perp\cap L_m} f(L_q+x+y)dydL_q\right)dxdL_m$$

$$= \int_{\mathcal{L}_m^n}\int_{\mathcal{L}_q^n(L_m)}\left(\int_{L_m^\perp}\int_{L_q^\perp\cap L_m} f(L_q+x+y)dydx\right)dL_qdL_m$$

$$= \int_{\mathcal{L}_m^n}\int_{\mathcal{L}_q^n(L_m)}\left(\int_{L_q^\perp} f(L_q+z)dz\right)dL_qdL_m$$

$$= \int_{\mathcal{L}_q^n}\int_{L_q^\perp} f(L_q+z)dzdL_q$$

$$= \int_{\mathcal{E}_q^n} f(E_q)dE_q,$$

where we have used the first integral formula and also Theorem 4.1.2. □

We remark that it is also possible to prove a uniqueness result for the measure μ_q, but the proof is a bit more involved. We also remark, that both measures ν_q and μ_q are actually independent of the choice of the reference space L_q^0.

Exercises and problems

1. Fill in the arguments omitted in the proof of Lemma 4.1.1 (invariance from the right and inversion invariance) by showing that

$$\nu(\mathcal{A}\vartheta) = \nu(\mathcal{A}), \qquad \text{and} \qquad \nu(\mathcal{A}^{-1}) = \nu(\mathcal{A}),$$

for all $\vartheta \in SO_n$ and all Borel sets $\mathcal{A} \subset SO_n$.

2. Show that

$$\int_{SO_n}\int_{(L_q^1)^\perp} f(\vartheta(L_q^1+x))dxd\nu(\vartheta) = \int_{SO_n}\int_{(L_q^0)^\perp} f(\vartheta(L_q^0+x))dxd\nu(\vartheta),$$

for $L_q^0, L_q^1 \in \mathcal{L}_q^n$ and a continuous function f on $\mathcal{E}_q^n$ with compact support (independence of the reference space).

* 3. Show that μ_q is the only invariant measure on $\mathcal{E}_q^n$ with

$$\mu_q(\mathcal{E}_q^n(B(1))) = \kappa_{n-q}.$$

4.2 Projection formulas

Theorem 4.2.1 (CAUCHY-KUBOTA). *Let $K \in \mathcal{K}^n$, $q \in \{0, \dots, n-1\}$ and $j \in \{0, \dots, q\}$. Then, we have*

$$\int_{\mathcal{L}_q^n} V_j(K|L_q)dL_q = \beta_{njq} V_j(K)$$

with

$$\beta_{njq} = \frac{\binom{q}{j}\kappa_q \kappa_{n-j}}{\binom{n}{j}\kappa_n \kappa_{q-j}}.$$

Proof. The mapping $L_q \mapsto K|L_q$ is continuous, therefore

$$L_q \mapsto V_j(K|L_q)$$

is continuous.

We first consider the case $q = n-1$. For $j = q$, we get CAUCHY's surface formula which has been proved already at the beginning of section 4.1. For $j < q$, we combine this with the STEINER formula (in dimension $n-1$). We obtain

$$\begin{aligned} V_{n-1}(K + B(\alpha)) &= \frac{n\kappa_n}{2\kappa_{n-1}} \int_{\mathcal{L}_{n-1}^n} V_{n-1}(K|L_{n-1} + [B(\alpha) \cap L_{n-1}])dL_{n-1} \\ &= \frac{n\kappa_n}{2\kappa_{n-1}} \sum_{j=0}^{n-1} \alpha^{n-1-j}\kappa_{n-1-j} \int_{\mathcal{L}_{n-1}^n} V_j(K|L_{n-1})dL_{n-1}. \end{aligned}$$

(Here, we make use of the fact that V_{n-1} is dimension invariant, hence $V_{n-1}(K|L_{n-1})$ yields the same value in L_{n-1} as in $\mathbb{R}^n$.) On the other hand, Corollary 3.5.4 (or Exercise 3.3.7) show that

$$\begin{aligned} V_{n-1}(K + B(\alpha)) &\left(= \frac{1}{2}F(K + B(\alpha))\right) \\ &= \sum_{j=0}^{n-1} \alpha^{n-1-j} \frac{(n-j)\kappa_{n-j}}{2} V_j(K). \end{aligned}$$

The formula for $j < q = n-1$ follows now by comparing the coefficients in these two polynomial expansions.

Finally, the case $q < n-1$ is obtained by a recursion. Namely, assume that the formula holds for $q+1$. Then, using Theorem 4.1.2 we obtain

$$\int_{\mathcal{L}_q^n} V_j(K|L_q)dL_q = \int_{\mathcal{L}_{q+1}^n} \left(\int_{\mathcal{L}_q^n(L_{q+1})} V_j(K|L_q)dL_q \right) dL_{q+1}.$$

The inner integral refers to the hyperplane case (in dimension $q+1$) which we have proved already. Therefore,

$$\begin{aligned}\int_{\mathcal{L}_q^n} V_j(K|L_q)dL_q &= \beta_{(q+1)jq}\int_{\mathcal{L}_{q+1}^n} V_j(K|L_{q+1})dL_{q+1}\\ &= \beta_{(q+1)jq}\beta_{nj(q+1)}V_j(K)\\ &= \beta_{njq}V_j(K),\end{aligned}$$

where we have used the assertion for $q+1$ and the fact that $K|L_q = (K|L_{q+1})|L_q$. □

Remarks. (1) For $j = q$, the CAUCHY-KUBOTA formulas yield

$$V_j(K) = \frac{1}{\beta_{njj}}\int_{\mathcal{L}_j^n} V_j(K|L_j)dL_j,$$

hence $V_j(K)$ is proportional to the mean content of the projections of K onto j-dimensional subspaces. Since $V_j(K|L_j)$ is also the content of the base of the cylinder circumscribed to K (with direction space $L^\perp$), $V_j(K|L_j)$ was called the 'quermass' of K in direction $L^\perp$. This explains the name 'quermassintegral' for the functionals $W_{n-j}(K) = c_{nj}V_j(K)$.
(2) For $j = q = 1$, we obtain

$$V_1(K) = \frac{1}{\beta_{n11}}\int_{\mathcal{L}_1^n} V_1(K|L_1)dL_1.$$

This gives now a rigorous proof for the fact that $V_1(K)$ is proportional to the mean width of K.

Exercises and problems

1. Prove the following generalizations of the CAUCHY-KUBOTA formulas:

(a)
$$\int_{\mathcal{L}_q^n} V^{(q)}(K_1|L_q,\dots,K_q|L_q)dL_q = \gamma_{nq}V(K_1,\dots,K_q,\underbrace{B(1),\dots,B(1)}_{n-q}),$$

for $K_1,\dots,K_q \in \mathcal{K}^n$, $0 \le q \le n-1$, and a certain constant γ_{nq},

(b)
$$\int_{\mathcal{L}_q^n} S_j^{(q)}(K|L_q, A\cap L_q)dL_q = \delta_{njq}S_j(K,A),$$

for $K \in \mathcal{K}^n$, a Borel set $A \subset S^{n-1}$, $0 \le j < q \le n-1$, and a certain constant δ_{njq}.

4.3 Section formulas

Theorem 4.3.1 (CROFTON). *Let $K \in \mathcal{K}^n$, $q \in \{0, \ldots, n-1\}$ and $j \in \{0, \ldots, q\}$. Then, we have*

$$\int_{\mathcal{E}_q^n} V_j(K \cap E_q) dE_q = \alpha_{njq} V_{n+j-q}(K)$$

with

$$\alpha_{njq} = \frac{\binom{q}{j} \kappa_q \kappa_{n+j-q}}{\binom{n}{q-j} \kappa_n \kappa_j}.$$

Proof. Here, we start with the case $j = 0$. From Theorem 4.1.3, we get

$$\int_{\mathcal{E}_q^n} V_0(K \cap E_q) dE_q = \int_{\mathcal{L}_q^n} \int_{L_q^\perp} V_0(K \cap (L_q + x)) dx dL_q.$$

On the right-hand side, the integrand fulfills

$$V_0(K \cap (L_q + x)) = \begin{cases} 1 & \text{if} \quad x \in K|L_q^\perp \,, \\ 0 & \phantom{\text{if}} \quad x \notin K|L_q^\perp \,. \end{cases}$$

Hence, using Theorems 4.1.2 and 4.2.1, we obtain

$$\begin{aligned} \int_{\mathcal{E}_q^n} V_0(K \cap E_q) dE_q &= \int_{\mathcal{L}_q^n} V_{n-q}(K|L_q^\perp) dL_q \\ &= \int_{\mathcal{L}_{n-q}^n} V_{n-q}(K|L_{n-q}) dL_{n-q} \\ &= \beta_{n(n-q)(n-q)} V_{n-q}(K) \\ &= \alpha_{n0q} V_{n-q}(K). \end{aligned}$$

This proves the result for $j = 0$.

Now, let $j > 0$. We use the result just proven for $K \cap E_q$ (in E_q) and obtain

$$V_j(K \cap E_q) = \frac{1}{\beta_{qjj}} \int_{\mathcal{E}_{q-j}^n(E_q)} V_0(K \cap E_{q-j}) dE_{q-j}.$$

Hence,

$$\begin{aligned} \int_{\mathcal{E}_q^n} V_j(K \cap E_q) dE_q &= \frac{1}{\beta_{qjj}} \int_{\mathcal{E}_q^n} \int_{\mathcal{E}_{q-j}^n(E_q)} V_0(K \cap E_{q-j}) dE_{q-j} dE_q \\ &= \frac{1}{\beta_{qjj}} \int_{\mathcal{E}_{q-j}^n} V_0(K \cap E_{q-j}) dE_{q-j} \\ &= \frac{\beta_{n(n+j-q)(n+j-q)}}{\beta_{qjj}} V_{n+j-q}(K) \\ &= \alpha_{njq} V_{n+j-q}(K), \end{aligned}$$

where we have first used Theorem 4.1.3 and then again the result above. □

Remarks. (1) Replacing the pair (j, q) by $(0, n-j)$, we obtain

$$\begin{aligned} V_j(K) &= \frac{1}{\alpha_{n0(n-j)}} \int_{\mathcal{E}_{n-j}^n} V_0(K \cap E_{n-j}) dE_{n-j} \\ &= \frac{1}{\alpha_{n0(n-j)}} \int_{K \cap E_{n-j} \neq \emptyset} dE_{n-j} \\ &= \frac{1}{\alpha_{n0(n-j)}} \mu_{n-j}(\{E_{n-j} \in \mathcal{E}_{n-j}^n : K \cap E_{n-j} \neq \emptyset\}). \end{aligned}$$

Hence, $V_j(K)$ is (up to a constant) the measure of all $(n-j)$-flats which meet K.
(2) We can give another interpretation of $V_j(K)$ in terms of flats touching K. Namely, consider the set

$$A(\alpha) := (\{E_{n-j-1} \in \mathcal{E}_{n-j-1}^n : K \cap E_{n-j-1} = \emptyset, K + B(\alpha) \cap E_{n-j-1} \neq \emptyset\}).$$

These are the $(n-j-1)$-flats meeting the parallel body $K + B(\alpha)$ but not K. If the limit

$$\lim_{\alpha \to 0} \frac{1}{\alpha} \mu_{n-j-1}(A(\alpha))$$

exists, we can interpret it as the measure of all $(n-j-1)$-flats touching K. Now (1) and Exercise 3.3.7 show that

$$\begin{aligned} \frac{1}{\alpha} \mu_{n-j-1}(A(\alpha)) &= \frac{\alpha_{n0(n-j-1)}}{\alpha} [V_{j+1}(K + B(\alpha)) - V_{j+1}(K)] \\ &= \frac{\alpha_{n0(n-j-1)}}{\alpha} \sum_{i=0}^{j} \alpha^{j+1-i} \binom{n-i}{n-j-1} \frac{\kappa_{n-i}}{\kappa_{n-j-1}} V_i(K) \\ &\to \alpha_{n0(n-j-1)} (n-j) \frac{\kappa_{n-j}}{\kappa_{n-j-1}} V_j(K), \end{aligned}$$

as $\alpha \to 0$.
(3) We can use (1) to solve some problems of *Geometrical Probability*. Namely, if $K, K_0 \in \mathcal{K}^n$ are such that $K \subset K_0$ and $V(K_0) > 0$, we can restrict μ_q to $\{E_q \in \mathcal{E}_q^n : K_0 \cap E_q \neq \emptyset\}$ and normalize it to get a probability measure. A random q-flat X_q with this distribution is called a *random q-flat in* K_0. We then get

$$\text{Prob}(X_q \cap K \neq \emptyset) = \frac{V_{n-q}(K)}{V_{n-q}(K_0)}.$$

As an example, we mention the BUFFON needle problem. Originally the problem was formulated in the following way: Given an array of parallel lines in the plane $\mathbb{R}^2$ with distance 1, what is the probability that a randomly thrown needle of length $L < 1$ intersects one of the lines? If we consider the disc of radius $\frac{1}{2}$ around the center of the needle, there will be almost surely exactly one line of the array intersecting this disc. Hence, the problem can be formulated in an equivalent way: Assume the needle N is fixed with center at 0. What is the probability that a random line X_1 in $B(\frac{1}{2})$ intersects N? The answer is

$$\begin{aligned}\mathrm{Prob}(X_1 \cap N \neq \emptyset) &= \frac{V_1(N)}{V_1(B(\frac{1}{2}))} \\ &= \frac{L}{\pi/2} \\ &= \frac{2L}{\pi}.\end{aligned}$$

(4) In continuation of (3), we can consider, for $K, K_0 \in \mathcal{K}^n$ with $K \subset K_0$ and $V(K_0) > 0$ and for a random q-flat X_q in K_0, the expected j-th intrinsic volume of $K \cap X_q$, $j \in \{0, \ldots, q\}$. We get

$$\begin{aligned}\mathbb{E}V_j(K \cap X_q) &= \frac{\int V_j(K \cap E_q) dE_q}{\int V_0(K_0 \cap E_q) dE_q} \\ &= \frac{\alpha_{njq} V_{n+j-q}(K)}{\alpha_{n0q} V_{n-q}(K_0)}.\end{aligned}$$

If K_0 is supposed to be known (and K is unknown) and if $V_j(K \cap X_q)$ is observable, then

$$\frac{\alpha_{n0q} V_{n-q}(K_0)}{\alpha_{njq}} V_j(K \cap X_q)$$

is an unbiased estimator of $V_{n+j-q}(K)$. Varying q, we get in this way three estimators for the volume $V(K)$, two for the surface area $F(K)$ and one for the mean width $\overline{B}(K)$.

The estimation formulas in Remark (4) would be of practical interest, if the set K under consideration was not supposed to be convex. In this final part, we therefore want to generalize the CROFTON formulas to certain non-convex sets. The set class which we consider is the *convex ring* $\mathcal{R}^n$, which consists of finite unions of convex bodies,

$$\mathcal{R}^n := \{\bigcup_{i=1}^{k} K_i : k \in \mathbb{N}, K_i \in \mathcal{K}^n\}.$$

We assume $\emptyset \in \mathcal{K}^n$, hence $\mathcal{R}^n$ is closed against finite unions and intersections, and it is the smallest set class with this property and containing $\mathcal{K}^n$. It is easy to see that $\mathcal{R}^n$ is dense in the class $\mathcal{C}^n$ of compact subsets of $\mathbb{R}^n$ (in the Hausdorff metric), hence any compact set can be well approximated by elements of $\mathcal{R}^n$.

Our first goal is to extend the intrinsic volumes V_j to sets in $\mathcal{R}^n$. Since V_j is additive on $\mathcal{K}^n$ (see the exercises), we seek an additive extension. Here a functional φ on $\mathcal{R}^n$ (or $\mathcal{K}^n$) is called *additive*, if

$$\varphi(K \cup M) + \varphi(K \cap M) = \varphi(K) + \varphi(M).$$

On $\mathcal{R}^n$, we require that this relation holds for all $K, M \in \mathcal{R}^n$, whereas on $\mathcal{K}^n$ we can only require it for $K, M \in \mathcal{K}^n$ with $K \cup M \in \mathcal{K}^n$. In addition, we assume that an additive functional φ fulfills $\varphi(\emptyset) = 0$. If $\varphi : \mathcal{R}^n \to \mathbb{R}$ is additive, the inclusion-exclusion principle (which follows by induction) shows that, for $A \in \mathcal{R}^n$, $A = \bigcup_{i=1}^k K_i$, $K_i \in \mathcal{K}^n$, we have

$$(*) \qquad \varphi(A) = \sum_{v \in S(k)} (-1)^{|v|-1} \varphi(K_v).$$

Here, we have used the following notation: $S(k)$ is the set of all non-empty finite subsets of $\{1, \dots, k\}$, $|v|$ is the cardinality of v, and K_v, for $v = \{i_1, \dots, i_m\}$, is the intersection $K_{i_1} \cap \dots \cap K_{i_m}$. $(*)$ shows that the values of φ on $\mathcal{R}^n$ depend only on the behavior of φ on $\mathcal{K}^n$. In particular, if an additive functional $\varphi : \mathcal{K}^n \to \mathbb{R}$ has an additive extension to $\mathcal{R}^n$, then this extension is unique. On the other hand, $(*)$ cannot be used to show the existence of such an additive extension, since the right-hand side may depend on the special representation $A = \bigcup_{i=1}^k K_i$ (and, in general, a set $A \in \mathcal{R}^n$ can have many different representations as a finite union of convex bodies). There is a general theorem of GROEMER which guarantees the existence of an additive extension for all functionals φ which are additive and continuous on $\mathcal{K}^n$. For the intrinsic volumes V_j, however, we can use a direct approach due to HADWIGER.

Theorem 4.3.2. *For $j = 0, \dots, n$, there is a unique additive extension of V_j onto $\mathcal{R}^n$.*

Proof. It remains to show the existence.

We begin with the Euler characteristic V_0 and prove the existence by induction on n, $n \geq 0$.

It is convenient to start with the dimension $n = 0$ since $\mathcal{R}^0 = \{\emptyset, \{0\}\} (= \mathcal{K}^0)$. Because of $V_0(\emptyset) = 0$ and $V_0(\{0\}) = 1$, V_0 is additive on $\mathcal{R}^0$.

For the step from dimension $n-1$ to dimension $n, n \geq 1$, we choose a fixed direction $u_0 \in S^{n-1}$ and consider the family of hyperplanes $E_\alpha := \{\langle \cdot, u_0 \rangle = \alpha\}, \alpha \in \mathbb{R}$. Then, for $A \in \mathcal{R}^n$, $A = \bigcup_{i=1}^k K_i$, $K_i \in \mathcal{K}^n$, we have

$$A \cap E_\alpha = \bigcup_{i=1}^k (K_i \cap E_\alpha)$$

and by induction hypothesis the additive extension $V_0(A \cap E_\alpha)$ exists. From $(*)$ we obtain that the function $f_A : \alpha \mapsto V_0(A \cap E_\alpha)$ is integer-valued and bounded from below and above. Therefore, f_A is piecewise constant and $(*)$ shows that the value of $f_A(\alpha)$ can only change if the hyperplane E_α supports one of the convex bodies K_v, $v \in S(k)$. We define the 'jump function'

$$g_A(\alpha) := f_A(\alpha) - \lim_{\beta \searrow \alpha} f_A(\beta), \qquad \alpha \in \mathbb{R},$$

and put

$$V_0(A) := \sum_{\alpha\in\mathbb{R}} g_A(\alpha).$$

This definition makes sense since $g_A(\alpha) \neq 0$ only for finitely many values of α. Moreover, for $k = 1$, that is $A = K \in \mathcal{K}^n$, $K \neq \emptyset$, we have $V_0(K) = 0 + 1 = 1$, hence V_0 is an extension of the Euler characteristic. By induction hypothesis, $A \mapsto f_A(\alpha)$ is additive on $\mathcal{R}^n$ for each α. Therefore, as a limit, $A \mapsto g_A(\alpha)$ is additive and so V_0 is additive. The uniqueness, which we have already obtained from $(*)$, shows that this construction does not depend on the choice of the direction u_0.

Now we consider the case $j > 0$. For $A \in \mathcal{R}^n$, $A = \bigcup_{i=1}^k K_i$, $K_i \in \mathcal{K}^n$, and $\alpha > 0$, $x \in \mathbb{R}^n$, we have

$$A \cap (B(\alpha) + x) = \bigcup_{i=1}^{k} (K_i \cap (B(\alpha) + x)).$$

Therefore, $(*)$ implies

$$V_0(A \cap (B(\alpha) + x)) = \sum_{v\in S(k)} (-1)^{|v|-1} V_0(K_v \cap (B(\alpha) + x)).$$

Since $V_0(K_v \cap (B(\alpha) + x)) = 1$, if and only if $x \in K_v + B(\alpha)$, we then get from the STEINER formula

$$\begin{aligned}
\int_{\mathbb{R}^n} V_0(A \cap (B(\alpha) + x))dx &= \sum_{v\in S(k)} (-1)^{|v|-1} \int_{\mathbb{R}^n} V_0(K_v \cap (B(\alpha) + x))dx \\
&= \sum_{v\in S(k)} (-1)^{|v|-1} V_n(K_v + B(\alpha)) \\
&= \sum_{v\in S(k)} (-1)^{|v|-1} \left(\sum_{j=0}^{n} \alpha^{n-j} \kappa_{n-j} V_j(K_v) \right) \\
&= \sum_{j=0}^{n} \alpha^{n-j} \kappa_{n-j} \left(\sum_{v\in S(k)} (-1)^{|v|-1} V_j(K_v) \right).
\end{aligned}$$

If we define

$$V_j(A) := \sum_{v\in S(k)} (-1)^{|v|-1} V_j(K_v),$$

then

$$\int_{\mathbb{R}^n} V_0(A \cap (B(\alpha) + x))dx = \sum_{j=0}^{n} \alpha^{n-j} \kappa_{n-j} V_j(A).$$

Since this equation holds for all $\alpha > 0$, the values $V_j(A), j = 0, \dots, n$, depend only on A and not on the special representation, and moreover V_j is additive. □

Remarks. (1) The formula

$$\int_{\mathbb{R}^n} V_0(A \cap (B(\alpha) + x))dx = \sum_{j=0}^{n} \alpha^{n-j} \kappa_{n-j} V_j(A),$$

which we used in the above proof, is a *generalized* STEINER *formula*; it reduces to the classical STEINER formula if $A \in \mathcal{K}^n$.

(2) The extended EULER characteristic V_0 (also called the EULER-POINCARE characteristic) plays also an important role in topology. In $\mathbb{R}^2$ and for $A \in \mathcal{R}^2$, $V_0(A)$ equals the number of connected components minus the number of 'holes' in A.

(3) On $\mathcal{R}^n$, V_n is still the volume (Lebesgue measure) and $F = 2V_{n-1}$ can still be interpreted as the surface area. The other (extended) intrinsic volumes V_j do not have a direct geometric interpretation.

Since union and intersection can be interchanged (as we have used already in the above arguments), the additivity of V_j allows us directly to extend the Crofton formulas to the convex ring.

Theorem 4.3.3 (CROFTON)**.** *Let $A \in \mathcal{R}^n$, $q \in \{0, \ldots, n-1\}$ and $j \in \{0, \ldots, q\}$. Then, we have*

$$\int_{\mathcal{E}_q^n} V_j(A \cap E_q) dE_q = \alpha_{njq} V_{n+j-q}(A).$$

As we have explained in a previous remark, these formulas can be used to give unbiased estimators for $V_{n+j-q}(A)$ based on intersections $A \cap X_q$ with random q-flats in the reference body K_0. This can be used in practical situations to estimate the surface area of a complicated tissue A in, say, a cubical specimen K_0 by measuring the boundary length $L(A \cap X_2)$ of a planar section $A \cap X_2$. Since the latter quantity is still complicated to obtain, one uses the CROFTON formulas again and estimates $L(A \cap X_2)$ from counting intersections with random lines X_1 in $K_0 \cap X_2$. Such *stereological formulas* are used and have been developed further in many applied sciences including medicine, biology, geology, metallurgy and material science.

Exercises and problems

1. Calculate the probability that a random secant of $B(1)$ is longer than $\sqrt{3}$. (According to the interpretation of a 'random secant', one might get here the values $\frac{1}{2}$, $\frac{1}{3}$ or $\frac{1}{4}$. Explain why $\frac{1}{2}$ is the right, 'rigid motion invariant' answer.)

2. Let $K, K' \in \mathcal{K}^n$ and $K \cup K' \in \mathcal{K}^n$. Show that:

   (a) $(K \cap K') + (K \cup K') = K + K'$,

   (b) $(K \cap K') + M = (K + M) \cap (K' + M)$, for all $M \in \mathcal{K}^n$.

   (c) $(K \cup K') + M = (K + M) \cup (K' + M)$, for all $M \in \mathcal{K}^n$.

3. Let $\varphi(K) := V(\underbrace{K, \ldots, K}_{j\text{-mal}}, M_{j+1}, \ldots, M_n)$, where $K, M_{j+1}, \ldots, M_n \in \mathcal{K}^n$. Show that φ is additive, that is
$$\varphi(K \cap K') + \varphi(K \cup K') = \varphi(K) + \varphi(K')$$
for all $K, K' \in \mathcal{K}^n$ with $K \cup K' \in \mathcal{K}^n$.

4. Show that the mappings $K \mapsto S_j(K, A)$ are additive on $\mathcal{K}^n$, for all $j \in \{0, \ldots, n\}$ and all Borel sets $A \subset S^{n-1}$.

5. Show that the convex ring $\mathcal{R}^n$ is dense in $\mathcal{C}^n$ in the HAUSDORFF metric.

6. Let $\varphi : \mathcal{R}^n \to \mathbb{R}$ be additive and $A \in \mathcal{R}^n, A = \bigcup_{i=1}^k K_i, K_i \in \mathcal{K}^n$. Give a proof for the inclusion-exclusion formula
$$(*) \qquad \varphi(A) = \sum_{v \in S(k)} (-1)^{|v|-1} \varphi(K_v).$$

Index