ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS 44

CONVEX BODIES: THE BRUNN-MINKOWSKI THEORY

ROLF SCHNEIDER

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Convex Bodies: The Brunn–Minkowski Theory

ROLF SCHNEIDER
University of Freiburg

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521352208

First published 1993
This digitally printed version 2008

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Schneider, Rolf, 1940–
Convex bodies: the Brunn–Minkowski theory / Rolf Schneider.
p. cm. – (Encyclopedia of mathematics and its applications; v. 44)
Includes bibliographical references and index.
ISBN 0-521-35220-7
1. Convex bodies. I. Title. II. Series.
QA649.S353 1993
516.3′74–dc20 92-11481 CIP

ISBN 978-0-521-35220-8 hardback
ISBN 978-0-521-05991-6 paperback

PREFACE

The Brunn–Minkowski theory is the classical core of the geometry of convex bodies. It originated with the thesis of Hermann Brunn in 1887 and is in its essential parts the creation of Hermann Minkowski, around the turn of the century. The well-known survey of Bonnesen and Fenchel in 1934 collected what was already an impressive body of results, though important developments were still to come, through the work of A.D. Aleksandrov and others in the thirties. In recent decades, the theory of convex bodies has expanded considerably; new topics have been developed and originally neglected branches of the subject have gained in interest. For instance, the combinatorial aspects, the theory of convex polytopes and the local theory of Banach spaces attract particular attention now. Nevertheless, the Brunn–Minkowski theory has remained of constant interest owing to its various new applications, its connections with other fields, and the challenge of some resistant open problems.

Aiming at a brief characterization of Brunn–Minkowski theory, one might say that it is the result of merging two elementary notions for point sets in Euclidean space: vector addition and volume. The vector addition of convex bodies, usually called Minkowski addition, has many facets of independent geometric interest. Combined with volume, it leads to the fundamental Brunn–Minkowski inequality and the notion of mixed volumes. The latter satisfy a series of inequalities which, due to their flexibility, solve many extremal problems and yield several uniqueness results. Looking at mixed volumes from a local point of view, one is led to mixed area measures. Quermassintegrals, or Minkowski functionals, and their local versions, surface area measures and curvature measures, are a special case of mixed volumes and mixed area measures. They are related to the differential geometry of convex hypersurfaces and to integral geometry.

Chapter 1 of the present book treats the basic properties of convex bodies and thus lays the foundations for subsequent developments. This chapter does not claim much originality; in large parts, it follows the procedures in standard books such as McMullen & Shephard [26], Roberts & Varberg [28], and Rockafellar [29]. Together with Sections 2.1, 2.2, 2.4, and 2.5, it serves as a general introduction to the metric geometry of convex bodies. Chapter 2 is devoted to the boundary structure of convex bodies. Most of its material is needed later, except for Section 2.6, on generic boundary structure, which just rounds off the picture. Minkowski addition is the subject of Chapter 3. Several different aspects are considered here such as decomposability, approximation problems with special regard to addition, additive maps and sums of segments. Quermassintegrals, which constitute a fundamental class of functionals on convex bodies, are studied in Chapter 4, where they are viewed as specializations of curvature measures, their local versions. For these, some integral-geometric formulae are established in Section 4.5. Here I try to follow the tradition set by Blaschke and Hadwiger of incorporating parts of integral geometry into the theory of convex bodies. Some of this, however, is also a necessary prerequisite for Section 4.6. The remaining part of the book is devoted to mixed volumes and their applications. Chapter 5 develops the basic properties of mixed volumes and mixed area measures and treats special formulae, extensions and analogues. Chapter 6, the heart of the book, is devoted to the inequalities satisfied by mixed volumes, with special emphasis on improvements, the equality cases (as far as they are known) and stability questions. Chapter 7 presents a small selection of applications. The classical theorems of Minkowski and the Aleksandrov–Fenchel–Jessen theorem are treated here, the latter in refined versions. Section 7.4 serves as an overview of affine extremal problems for convex bodies. In this promising field, Brunn–Minkowski theory is of some use, but it appears that for the solution of some long-standing open problems new methods still have to be invented.

Concerning the choice of topics treated in this book, I wish to point out that it is guided by Minkowski's original work also in the following sense. Some subjects that Minkowski touched only briefly have later expanded considerably, and I pay special attention to these. Examples are projection bodies (zonoids), tangential bodies, the use of spherical harmonics in convexity and strengthenings of Minkowskian inequalities in the form of stability estimates.

The necessary prerequisites for reading this book are modest: the usual geometry of Euclidean space, elementary analysis, and basic measure and integration theory for Chapter 4. Occasionally, use is made of spherical harmonics; relevant information is collected in the Appendix. My intended attitude towards the presentation of proofs cannot be summarized better than by quoting from the preface to the book on Hausdorff measures by C.A. Rogers: 'As the book is largely based on lectures, and as I like my students to follow my lectures, proofs are given in great detail; this may bore the mature mathematician, but it will I believe, be a great help to anyone trying to learn the subject *ab initio*.' On the other hand, some important results are stated as theorems but not proved, since this would lead us too far from the main theme, and no proofs are given in the survey sections 5.4, 6.8, and 7.4.

The notes at the end of nearly all sections are an essential part of the book. As a rule, this is where I have given references to original literature, considered questions of priority, made various comments and, in particular, given hints about applications, generalizations and ramifications. As an important purpose of the notes is to demonstrate the connections of convex geometry with other fields, some notes do take us further from the main theme of the book, mentioning, for example, infinite-dimensional results or non-convex sets or giving more detailed information on applications in, for instance, stochastic geometry.

The list of references does not have much overlap with the older bibliographies in the books by Bonnesen & Fenchel and by Hadwiger. Hence, a reader wishing to have a more complete picture should consult these bibliographies also, as well as those in the survey articles listed in part B of the References.

My thanks go to Sabine Linsenbold for her careful typing of the manuscript and to Daniel Hug who read the typescript and made many valuable comments and suggestions.

CONVENTIONS AND NOTATION

Here we shall fix our notation and collect some basic definitions. We shall work in n-dimensional real Euclidean vector space, $\mathbb{E}^n$, with origin o, scalar product $\langle \cdot, \cdot \rangle$ and induced norm $|\cdot|$. We shall not distinguish formally between the vector space $\mathbb{E}^n$ and its corresponding affine space, although our alternating use of the words 'vector' and 'point' is deliberate and should support the reader's intuition. As a rule, elements of $\mathbb{E}^n$ are denoted by lower-case letters, subsets by capitals and real numbers by small Greek letters. However, in later chapters the reader will notice an increasing number of exceptions to this rule.

The vector $x \in \mathbb{E}^n$ is a *linear combination* of the vectors $x_1, \ldots, x_k \in \mathbb{E}^n$ if $x = \lambda_1 x_1 + \ldots + \lambda_k x_k$ with suitable $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. If such λ_i exist with $\lambda_1 + \ldots + \lambda_k = 1$, then x is an *affine combination* of $x_1, \ldots, x_k$. For $A \subset \mathbb{E}^n$, $\operatorname{lin} A$ ($\operatorname{aff} A$) denotes the *linear hull (affine hull)* of A; this is the set of all linear (affine) combinations of elements of A and at the same time the smallest linear subspace (affine subspace) of $\mathbb{E}^n$ containing A. Points $x_1, \ldots, x_k \in \mathbb{E}^n$ are *affinely independent* if none of them is an affine combination of the others, i.e., if

$$\sum_{i=1}^{k} \lambda_i x_i = o \quad \text{with } \lambda_i \in \mathbb{R} \text{ and } \sum_{i=1}^{k} \lambda_i = 0$$

implies that $\lambda_1 = \ldots = \lambda_k = 0$. This is equivalent to the linear independence of the vectors $x_2 - x_1, \ldots, x_k - x_1$. We may also define a map $\tau: \mathbb{E}^n \to \mathbb{E}^n \times \mathbb{R}$ by $\tau(x) := (x, 1)$; then $x_1, \ldots, x_k \in \mathbb{E}^n$ are affinely independent if and only if $\tau(x_1), \ldots, \tau(x_k)$ are linearly independent.

For $x, y \in \mathbb{E}^n$ we write

$$[x, y] := \{(1 - \lambda)x + \lambda y \mid 0 \leqq \lambda \leqq 1\}$$

for the *closed segment* and

$$[x, y) := \{(1 - \lambda)x + \lambda y \mid 0 \leqq \lambda < 1\}$$

for a *half-open segment*, both with endpoints x, y. For $A, B \subset \mathbb{E}^n$ and $\lambda \in \mathbb{R}$ we define

$$A + B := \{a + b | a \in A, b \in B\},$$
$$\lambda A := \{\lambda a | a \in A\},$$

and we write $-A$ for $(-1)A$, $A - B$ for $A + (-B)$ and $A + x$ for $A + \{x\}$, where $x \in \mathbb{E}^n$. The set $A + B$ is written $A \oplus B$ and called the *direct sum* of A and B if A and B are contained in complementary affine subspaces of $\mathbb{E}^n$.

By $\operatorname{cl} A$, $\operatorname{int} A$, $\operatorname{bd} A$ we denote, respectively, the closure, interior and boundary of a subset A of a topological space. For $A \subset \mathbb{E}^n$, the sets $\operatorname{relint} A$, $\operatorname{relbd} A$ are the relative interior and relative boundary, that is, the interior and boundary of A relative to its affine hull.

The scalar product in $\mathbb{E}^n$ will often be used to describe hyperplanes and halfspaces. A *hyperplane* of $\mathbb{E}^n$ can be written in the form

$$H_{u,\alpha} = \{x \in \mathbb{E}^n | \langle x, u \rangle = \alpha\}$$

with $u \in \mathbb{E}^n \backslash \{o\}$ and $\alpha \in \mathbb{R}$; here $H_{u,\alpha} = H_{v,\beta}$ if and only if $(v, \beta) = (\lambda u, \lambda\alpha)$ with $\lambda \neq 0$. We say that u is a *normal vector* of $H_{u,\alpha}$. The hyperplane $H_{u,\alpha}$ bounds the two *closed halfspaces*

$$H^-_{u,\alpha} := \{x \in \mathbb{E}^n | \langle x, u \rangle \leqq \alpha\},$$
$$H^+_{u,\alpha} := \{x \in \mathbb{E}^n | \langle x, u \rangle \geqq \alpha\}.$$

Occasionally we also use $\langle \cdot , \cdot \rangle$ to denote the scalar product on $\mathbb{E}^n \times \mathbb{R}$ given by

$$\langle (x, \xi), (y, \eta) \rangle = \langle x, y \rangle + \xi\eta.$$

An affine subspace of $\mathbb{E}^n$ is often called a *flat*, and the intersection of a flat with a closed halfspace meeting the flat but not entirely containing it will be called a *half-flat*. A one-dimensional flat is a *line* and a one-dimensional half-flat a *ray*.

The following metric notions will be used. For $x, y \in \mathbb{E}^n$ and $\emptyset \neq A \subset \mathbb{E}^n$, $|x - y|$ is the *distance* between x and y and

$$d(A, x) := \inf \{|x - a| \,|\, a \in A\}$$

is the distance of x from A. For a bounded set $\emptyset \neq A \subset \mathbb{E}^n$,

$$\operatorname{diam} A := \sup \{|x - y| \,|\, x, y \in A\}$$

is the *diameter* of A. We write

$$B(z, \rho) := \{x \in \mathbb{E}^n \,|\, |x - z| \leqq \rho\}$$

and

$$B_0(z, \rho) := \{x \in \mathbb{E}^n \,|\, |x - z| < \rho\}$$

respectively for the closed and open balls with centre $z \in \mathbb{E}^n$ and radius

$\rho > 0$. $B^n := B(o, 1)$ is the *unit ball* and

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

the *unit sphere* of $\mathbb{E}^n$.

By $\mathcal{H}^k$ we denote the k-dimensional Hausdorff (outer) measure on $\mathbb{E}^n$, where $0 \leqq k \leqq n$. If A is a Borel subset of a k-dimensional flat E^k or a k-dimensional sphere S^k in $\mathbb{E}^n$, then $\mathcal{H}^k(A)$ coincides respectively with the k-dimensional Lebesgue measure of A computed in E^k or with the k-dimensional spherical Lebesgue measure of A computed in S^k. Hence, all integrations with respect to these Lebesgue measures can be expressed by means of the Hausdorff measure $\mathcal{H}^k$. In integrals with respect to $\mathcal{H}^n$ we often abbreviate $d\mathcal{H}^n(x)$ by dx. The n-dimensional measure of the unit ball in $\mathbb{E}^n$ is denoted by κ_n, and its surface area by ω_n, thus

$$\kappa_n = \mathcal{H}^n(B^n) = \frac{\pi^{n/2}}{\Gamma\left(1 + \frac{n}{2}\right)}, \quad \omega_n = \mathcal{H}^{n-1}(S^{n-1}) = n\kappa_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}.$$

Linear maps, affine maps and isometries between Euclidean spaces are defined as usual. In particular, a map $\varphi: \mathbb{E}^n \to \mathbb{E}^n$ is a *translation* if $\varphi(x) = x + t$ for $x \in \mathbb{E}^n$ with some fixed vector $t \in \mathbb{E}^n$, the *translation vector*. The set $A + t$ is called the *translate* of A by t. The map φ is a *homothety* if $\varphi(x) = \lambda x + t$ for $x \in \mathbb{E}^n$ with some $\lambda > 0$ and some $t \in \mathbb{E}^n$. The set $\lambda A + t$ with $\lambda > 0$ is called a *homothet* of A. Sets A, B are called *positively homothetic* if $A = \lambda B + t$ with $t \in \mathbb{E}^n$ and $\lambda > 0$, and *homothetic* if either they are positively homothetic or one of them is a singleton (a one-pointed set). A *rigid motion* of $\mathbb{E}^n$ is an isometry of $\mathbb{E}^n$ onto itself, and it is a *rotation* if it is an isometry fixing the origin. Each rigid motion is the composition of a rotation and a translation. A rigid motion is called *proper* if it preserves the orientation of $\mathbb{E}^n$, otherwise it is called *improper*. A rotation of $\mathbb{E}^n$ is a linear map; it preserves the scalar product and can be represented, with respect to an orthonormal basis, by an orthogonal matrix; this matrix has determinant 1 if and only if the rotation is proper. The composition of a rigid motion and a *dilatation*, by which we mean a map $x \mapsto \lambda x$ with $\lambda > 0$, is called a *similarity*.

By $SO(n)$ we denote the group of proper rotations of $\mathbb{E}^n$. With the topology induced by the usual matrix norm it is a compact topological group. The group of proper rigid motions of $\mathbb{E}^n$ is denoted by G_n and topologized as usual. Also, the Grassmannian $G(n, k)$ of k-dimensional linear subspaces of $\mathbb{E}^n$ and the set $A(n, k)$ of k-dimensional affine subspaces of $\mathbb{E}^n$ are endowed with their standard topologies.

The Haar measures on $SO(n)$, G_n, $A(n, k)$ are denoted respectively by ν, μ, μ_k. We normalize ν by $\nu(SO(n)) = 1$. The normalizations of the measures μ and μ_k will be fixed in Section 4.5 when they are needed.

For an affine subspace E of $\mathbb{E}^n$ we denote by proj_E the orthogonal projection from $\mathbb{E}^n$ onto E. We often write $\text{proj}_E A =: A|E$ for $A \subset \mathbb{E}^n$ (since A is a set, no confusion with the restriction of a function, for example $f|E$, can arise).

Some final remarks are in order. Since any k-dimensional affine subspace E of $\mathbb{E}^n$ is the image of $\mathbb{E}^k$ under some isometry, it is clear (and common practice without mention) that all notions and results that have been established for $\mathbb{E}^k$ and are invariant under isometries can be applied in E; similarly for affine-invariant notions and results.

The following notational conventions will be useful in several places. If f is a homogeneous function on $\mathbb{E}^n$, then $\bar{f}$ denotes its restriction to the unit sphere S^{n-1}. Very often, mappings of the type $f: \mathcal{K} \times \mathbb{E}^n \to M$ will occur where $\mathcal{K}$ is some class of subsets of $\mathbb{E}^n$. In this case we usually abbreviate, for fixed $K \in \mathcal{K}$, the function $f(K, \cdot): \mathbb{E}^n \to M$ by f_K.

Finally we wish to point out that in definitions the word 'if' is always understood as 'if and only if'.

CONTENTS

1

Basic convexity

1.1. Convex sets and combinations

A set $A \subset \mathbb{E}$ is *convex* if together with any two points x, y it contains the segment $[x, y]$, thus if

$$(1-\lambda)x + \lambda y \in A \quad \text{for } x, y \in A \quad \text{and } 0 \leqq \lambda \leqq 1.$$

Examples of convex sets are obvious; but observe also that $B_0(z, \rho) \cup A$ is convex if A is an arbitrary subset of the boundary of the open ball $B_0(z, \rho)$. As immediate consequences of the definition we note that intersections of convex sets are convex, affine images and pre-images of convex sets are convex and if A, B are convex, then $A + B$ and λA ($\lambda \in \mathbb{R}$) are convex.

Remark 1.1.1. For $A \subset \mathbb{E}^n$ and λ, $\mu > 0$ one trivially has $\lambda A + \mu A \supset (\lambda + \mu)A$. Equality (for all λ, $\mu > 0$) holds precisely if A is convex. In fact, if A is convex and $x \in \lambda A + \mu A$, then $x = \lambda a + \mu b$ with a, $b \in A$ and hence

$$x = (\lambda + \mu)\left(\frac{\lambda}{\lambda + \mu} a + \frac{\mu}{\lambda + \mu} b\right) \in (\lambda + \mu)A;$$

thus $\lambda A + \mu A = (\lambda + \mu)A$. The other direction of the assertion is trivial.

A set $A \subset \mathbb{E}^n$ is called a *convex cone* if A is convex and nonempty and if $x \in A$, $\lambda \geqq 0$ implies $\lambda x \in A$. Thus a nonempty set $A \subset \mathbb{E}^n$ is a convex cone if and only if A is closed under addition and under multiplication by non-negative real numbers.

By restricting affine and linear combinations to non-negative coefficients, one obtains the following two fundamental notions. The point $x \in \mathbb{E}^n$ is a *convex combination* of the points $x_1, \ldots, x_k \in \mathbb{E}^n$ if there are numbers $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that

$$x = \sum_{i=1}^{k} \lambda_i x_i, \ \lambda_i \geqq 0 \quad (i = 1, \ldots, k), \ \sum_{i=1}^{k} \lambda_i = 1.$$

Similarly, the vector $x \in \mathbb{E}^n$ is a *positive combination* of the vectors $x_1, \ldots, x_k \in \mathbb{E}^n$ if

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{with } \lambda_i \geqq 0 \quad (i = 1, \ldots, k).$$

For $A \subset \mathbb{E}^n$ the set of all convex combinations (positive combinations) of any finitely many elements of A is called the *convex hull* (*positive hull*) of A and is denoted by $\operatorname{conv} A$ ($\operatorname{pos} A$).

Theorem 1.1.2. *If $A \subset \mathbb{E}^n$ is convex, then* $\operatorname{conv} A = A$. *For an arbitrary set $A \subset \mathbb{E}^n$,* $\operatorname{conv} A$ *is the intersection of all convex subsets of $\mathbb{E}^n$ containing A. If A, $B \subset \mathbb{E}^n$, then* $\operatorname{conv}(A + B) = \operatorname{conv} A + \operatorname{conv} B$.

Proof. Let A be convex. Trivially, $A \subset \operatorname{conv} A$. By induction we show that A contains all convex combinations of any k points of A. For $k = 2$ this holds by the definition of convexity. Suppose that it holds for $k - 1$ and that $x = \lambda_1 x_1 + \ldots + \lambda_k x_k$ with $x_1, \ldots, x_k \in A$, $\lambda_1 + \ldots + \lambda_k = 1$ and $\lambda_1, \ldots, \lambda_k > 0$, without loss of generality. Then

$$x = (1 - \lambda_k) \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i + \lambda_k x_k \in A$$

since

$$\frac{\lambda_i}{1 - \lambda_k} > 0, \quad \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} = 1$$

and hence

$$\sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i \in A$$

by hypothesis. This proves $A = \operatorname{conv} A$. For arbitrary $A \subset \mathbb{E}^n$ let $D(A)$ be the intersection of all convex sets $K \subset \mathbb{E}^n$ containing A. Since $A \subset \operatorname{conv} A$ and $\operatorname{conv} A$ is evidently convex, we have $D(A) \subset \operatorname{conv} A$. Each convex K with $A \subset K$ satisfies $\operatorname{conv} A \subset \operatorname{conv} K = K$, hence $\operatorname{conv} A \subset D(A)$, which proves the equality.

Let A, $B \subset \mathbb{E}^n$. Let $x \in \operatorname{conv}(A + B)$, thus

$$x = \sum_{i=1}^{k} \lambda_i (a_i + b_i) \quad \text{with } a_i \in A, \quad b_i \in B, \quad \lambda_i \geqq 0, \quad \sum_{i=1}^{k} \lambda_i = 1$$

and hence $x = \sum \lambda_i a_i + \sum \lambda_i b_i \in \operatorname{conv} A + \operatorname{conv} B$. Let $x \in \operatorname{conv} A + \operatorname{conv} B$, thus

$$x = \sum_i \lambda_i a_i + \sum_j \mu_j b_j$$

with $a_i \in A$, $b_j \in B$, $\lambda_i, \mu_j \geqq 0$, $\sum \lambda_i = \sum \mu_j = 1$. We may write

$$x = \sum_{i,j} \lambda_i \mu_j (a_i + b_j)$$

and deduce that $x \in \operatorname{conv}(A + B)$. ■

An immediate consequence is that $\operatorname{conv}(\operatorname{conv} A) = \operatorname{conv} A$.

Theorem 1.1.3. *If* $A \subset \mathbb{E}^n$ *is a convex cone, then* $\operatorname{pos} A = A$. *For a nonempty set* $A \subset \mathbb{E}^n$, $\operatorname{pos} A$ *is the intersection of all convex cones in* $\mathbb{E}^n$ *containing* A. *If* $A, B \subset \mathbb{E}^n$, *then* $\operatorname{pos}(A + B) = \operatorname{pos} A + \operatorname{pos} B$.

Proof. As above. ■

The following result on the generation of convex hulls is fundamental.

Theorem 1.1.4 (Carathéodory's theorem). *If* $A \subset \mathbb{E}^n$ *and* $x \in \operatorname{conv} A$, *then* x *is a convex combination of affinely independent points of* A. *In particular,* x *is a convex combination of* $n + 1$ *or fewer points of* A.

Proof. The point $x \in \operatorname{conv} A$ has a representation

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \text{with } x_i \in A, \quad \lambda_i > 0, \quad \sum_{i=1}^{k} \lambda_i = 1$$

with some $k \in \mathbb{N}$, and we may assume that k is minimal. Suppose that $x_1, \ldots, x_k$ are affinely dependent. Then there are numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, not all zero, with

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

We can choose m such that λ_m/α_m is positive and, with this restriction, as small as possible (observe that all λ_i are positive and at least one α_i is positive). In the affine representation

$$x = \sum_{i=1}^{k} \left(\lambda_i - \frac{\lambda_m}{\alpha_m} \alpha_i \right) x_i$$

all coefficients are non-negative (trivially, if $\alpha_i \leqq 0$, otherwise by the choice of m) and at least one of them is zero. This contradicts the minimality of k. Thus $x_1, \ldots, x_k$ are affinely independent, which implies $k \leqq n + 1$. ■

The convex hull of finitely many points is called a *polytope*. A *k-simplex* is the convex hull of $k + 1$ affinely independent points, and these points are the *vertices* of the simplex. Thus Carathéodory's theorem states that conv A is the union of all simplices with vertices in A.

Another equally simple and important result on convex hulls is the following.

Theorem 1.1.5 (Radon's theorem). *Each set of affinely dependent points (in particular, each set of at least $n + 2$ points) in $\mathbb{E}^n$ can be expressed as the union of two disjoint sets whose convex hulls have a common point.*

Proof. If $x_1, \ldots, x_k$ are affinely dependent, there are numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, not all zero, with

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

We may assume, after renumbering, that $\alpha_i > 0$ precisely for $i = 1, \ldots, j$; then $1 \leqq j < k$ (at least one α_i is $\neq 0$, say > 0, but not all α_i are > 0). With

$$\alpha := \alpha_1 + \ldots + \alpha_j = -(\alpha_{j+1} + \ldots + \alpha_k) > 0$$

we obtain

$$x := \sum_{i=1}^{j} \frac{\alpha_i}{\alpha} x_i = \sum_{i=j+1}^{k} \left(-\frac{\alpha_i}{\alpha}\right) x_i$$

and thus $x \in \operatorname{conv}\{x_1, \ldots, x_j) \cap \operatorname{conv}\{x_{j+1}, \ldots, x_k\}$. The assertion follows. ■

From Radon's theorem one easily deduces Helly's theorem, a fundamental and typical result of the combinatorial geometry of convex sets.

Theorem 1.1.6 (Helly's theorem). *Let $A_1, \ldots, A_k \subset \mathbb{E}^n$ be convex sets. If any $n + 1$ of these sets have a common point, then all the sets have a common point.*

Proof. Suppose that $k > n + 1$ (for $k < n + 1$ there is nothing to prove, and for $k = n + 1$ the assertion is trivial) and that the assertion is proved for $k - 1$ convex sets. Then for $i \in \{1, \ldots, k\}$ there exists a point

$$x_i \in A_1 \cap \ldots \cap \check{A}_i \cap \ldots \cap A_k$$

where $\check{A}_i$ indicates that A_i has been deleted. The $k \geqq n + 2$ points $x_1, \ldots, x_k$ are affinely dependent; hence from Radon's theorem we can infer that, after renumbering, there is a point

$$x \in \operatorname{conv}\{x_1, \ldots, x_j\} \cap \operatorname{conv}\{x_{j+1}, \ldots, x_k\}$$

for some $j \in \{1, \ldots, k-1\}$. Because $x_1, \ldots, x_j \in A_{j+1}, \ldots, A_k$ we have

$$x \in \operatorname{conv}\{x_1, \ldots, x_j\} \subset A_{j+1} \cap \ldots \cap A_k,$$

similarly $x \in \operatorname{conv}\{x_{j+1}, \ldots, x_k\} \subset A_1 \cap \ldots \cap A_j$. ■

Here is a little example (another one is Theorem 1.3.11) to demonstrate how Helly's theorem can be applied to obtain elegant results of a similar nature:

Theorem 1.1.7. *Let $\mathcal{M}$ be a finite family of convex sets in $\mathbb{E}^n$ and let $K \subset \mathbb{E}^n$ be convex. If any $n+1$ elements of $\mathcal{M}$ are intersected by some translate of K, then all elements of $\mathcal{M}$ are intersected by a translate of K.*

Proof. Let $\mathcal{M} = \{A_1, \ldots, A_k\}$. To any $n+1$ elements of $\{1, \ldots, k\}$, say $1, \ldots, n+1$, there are $t \in \mathbb{E}^n$ and $x_i \in A_i \cap (K+t)$, hence $-t \in K - A_i$, for $i = 1, \ldots, n+1$. Thus any $n+1$ elements of the family $\{K - A_1, \ldots, K - A_k\}$ have nonempty intersection. By Helly's theorem there is a vector $-t \in \mathbb{E}^n$ with $-t \in K - A_i$ and hence $A_i \cap (K+t) \neq \emptyset$ for $i \in \{1, \ldots, k\}$. ■

Next we look at the interplay between convexity and topological properties. We start with a simple observation.

Lemma 1.1.8. *Let $A \subset \mathbb{E}^n$ be convex. If $x \in \operatorname{int} A$ and $y \in \operatorname{cl} A$, then $[x, y) \subset \operatorname{int} A$.*

Proof. Let $z = (1-\lambda)y + \lambda x$ with $0 < \lambda < 1$. We have $B(x, \rho) \subset A$ for some $\rho > 0$; put $B(o, \rho) =: U$. First we assume $y \in A$. Let $w \in \lambda U + z$, hence $w = \lambda u + z$ with $u \in U$. Then $x + u \in A$, hence $w = (1-\lambda)y + \lambda(x+u) \in A$. This shows that $\lambda U + z \subset A$ and thus $z \in \operatorname{int} A$.

Now assume merely that $y \in \operatorname{cl} A$. Put $V := [\lambda/(1-\lambda)]U + y$. There is some $a \in A \cap V$. We have $a = [\lambda/(1-\lambda)]u + y$ with $u \in U$ and hence $z = (1-\lambda)a + \lambda(x-u) \in A$. This proves that $[x, y) \subset A$, which together with the first part yields $[x, y) \subset \operatorname{int} A$. ■

Theorem 1.1.9. *If $A \subset \mathbb{E}^n$ is convex, then $\operatorname{int} A$ and $\operatorname{cl} A$ are convex. If $A \subset \mathbb{E}^n$ is open, then $\operatorname{conv} A$ is open.*

Proof. The convexity of $\operatorname{int} A$ follows from Lemma 1.1.8. The convexity of $\operatorname{cl} A$ for convex A and the openness of $\operatorname{conv} A$ for open A are easy exercises. ■