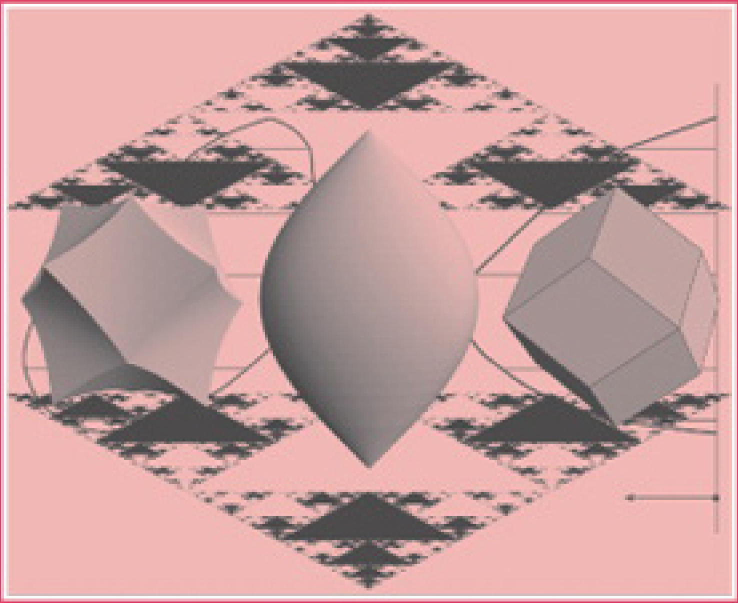
Encyclopedia of Mathematics and Its Applications 58

GEOMETRIC TOMOGRAPHY

SECOND EDITION

Richard J. Gardner



CAMBRIDGE

Geometric Tomography, second edition

Geometric tomography deals with the retrieval of information about a geometric object from data concerning its projections (shadows) on planes or cross-sections by planes. It is a geometric relative of computerized tomography, which reconstructs an image from X-rays of a human patient. The subject overlaps with convex geometry and employs many tools from that area, including some formulas from integral geometry. It also has connections to discrete tomography, geometric probing in robotics, and stereology.

This comprehensive study provides a rigorous treatment of the subject. Although primarily meant for researchers and graduate students in geometry and tomography, brief introductions, suitable for advanced undergraduates, are provided to the basic concepts. More than 70 illustrations are used to clarify the text. The book also presents 66 unsolved problems. Each chapter ends with extensive notes, historical remarks, and some biographies. This new edition includes numerous updates and improvements, with some 50 extra pages of material and 300 new references, bringing the total to more than 800.

Richard J. Gardner has been Professor of Mathematics at Western Washington University since 1991. He is the author of 70 papers and founded geometric tomography as a subject in its own right with the publication of the first edition of this book in 1995.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

58 R. J. Gardner *Geometric Tomography, 2ed.*
59 G. A. Baker Jr. and P. Graves-Morris *Padé Approximants, 2ed.*
60 J. Krajíček *Bounded Arithmetic Propositional Logic and Complexity Theory*
61 H. Groemer *Geometric Applications of Fourier Series and Spherical Harmonics*
62 H. O. Fattorini *Infinite Dimensional Optimization and Control Theory*
63 A. C. Thompson *Minkowski Geometry*
64 R. B. Bapat and T. E. S. Raghavan *Nonnegative Matrices and Applications*
65 K. Engel *Sperner Theory*
66 D. Cvetkovič, P. Rowlinson, and S. Simič *Eigenspaces of Graphs*
67 F. Bergeron, G. Labelle, and P. Leroux *Combinatorial Species and Tree-Like Structures*
68 R. Goodman and N. R. Wallach *Representations and Invariants of the Classical Groups*
69 T. Beth, D. Jungnickel, and H. Lenz *Design Theory I, 2ed.*
70 A. Pietsch and J. Wenzel *Orthonormal Systems and Banach Space Geometry*
71 G. E. Andrews, R. Askey, and R. Roy *Special Functions*
72 R. Ticciati *Quantum Field Theory for Mathematicians*
73 M. Stern *Semimodular Lattices*
74 I. Lasiecka and R. Triggiani *Control Theory for Partial Differential Equations I*
75 I. Lasiecka and R. Triggiani *Control Theory for Partial Differential Equations II*
76 A. A. Ivanov *Geometry of Sporadic Groups I*
77 A. Schinzel *Polynomials with Special Regard to Reducibility*
78 T. Beth, D. Jungnickel, and H. Lenz *Design Theory II, 2ed.*
79 T. W. Palmer *Banach Algebras and the General Theory of *-Algebras II*
80 O. Stormark *Lie's Structural Approach to PDE Systems*
81 C. F. Dunkl and Y. Xu *Orthogonal Polynomials of Several Variables*
82 J. P. Mayberry *The Foundations of Mathematics in the Theory of Sets*
83 C. Foias, O. Manley, R. M. S. Rosa, and R. Temam *Navier-Stokes Equations and Turbulence*
84 B. Polster and G. Steinke *Geometries on Surfaces*
85 R. B. Paris and D. Kaminski *Asymptotics and Mellin-Barnes Integrals*
86 R. J. McEliece *The Theory of Information and Coding, 2ed.*
87 B. A. Magurn *An Algebraic Introduction to K-Theory*
88 T. Mora *Solving Polynomial Equation Systems I*
89 K. Bichteler *Stochastic Integration with Jumps*
90 M. Lothaire *Algebraic Combinatorics on Words*
91 A. A. Ivanov and S. V. Shpectorov *Geometry of Sporadic Groups, 2ed.*
92 P. McMullen and E. Schulte *Abstract Regular Polytopes*
93 G. Gierz *et al.* *Continuous Lattices and Domains*
94 S. R. Finch *Mathematical Constants*
95 Y. Jabri *The Mountain Pass Theorem*
96 G. Gasper and M. Rahman *Basic Hypergeometric Series, 2ed.*
97 M. C. Pedicchio and W. Tholen *Categorical Foundations*
98 M. Ismail *Classical and Quantum Orthogonal Polynomials in One Variable*
99 T. Mora *Solving Polynomial Equation Systems II*
100 E. Olivieri and M. E. Vares *Large Deviations and Metastability*
101 A. Kushner, V. Lychagin, and V. Rubtsov *Contact Geometry and Nonlinear Differential Equations*
102 L. W. Beineke and R. J. Wilson *Topics in Algebraic Graph Theory*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Geometric Tomography

Second Edition

RICHARD J. GARDNER
Western Washington University

CAMBRIDGE
UNIVERSITY PRESS

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

Cambridge University Press
40 West 20th Street, New York, NY 10011-4211, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521866804

First edition published 1995
Second edition published 2006

Printed in the United States of America

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Gardner, Richard J.
Geometric tomography / Richard J. Gardner.– 2nd ed.
p. cm. – (Encyclopedia of mathematics and its applications ; v. 58)
Includes bibliographical references and index.
ISBN-13: 978-0-521-86680-4 (hardback)
ISBN-10: 0-521-86680-4 (hardback)
ISBN-13: 978-0-521-68493-4 (pbk.)
ISBN-10: 0-521-68493-5 (pbk.)
1. Geometric tomography. I. Title. II. Series.
QA639.5.G37 2006
516.3′62 – dc22 2006002945

ISBN-13 978-0-521-86680-4 hardback
ISBN-10 0-521-86680-4 hardback

ISBN-13 978-0-521-68493-4 paperback
ISBN-10 0-521-68493-5 paperback

To Linda

But one man loved the pilgrim soul in you

CONTENTS

PREFACE TO THE SECOND EDITION

This second edition incorporates some 60 extra pages of material, including seven new figures, another 21 chapter notes, the new Sections 4.4 and A.4, and about 300 additional references. This expansion indicates the amazingly rapid development of geometric tomography over the decade since the first edition appeared. Despite this, the list of 66 open problems is of roughly the same length.

Many corrections have been made. The most significant amendment appears in Chapter 8, written for the first edition very shortly after the pioneering work was published. Alex Koldobsky's work described in Note 8.9 brought to light an error in the solution of the Busemann–Petty problem. The revised Chapter 8 contains the corrected solution, while the six new notes for Chapter 8 struggle to keep pace with the incredible activity around intersection bodies. The task of surveying recent developments would have been a great deal more difficult but for the publication of Koldobsky's fine book [465]. This takes an almost entirely analytic point of view, whereas in Chapter 8 the original geometrical approach is retained as far as possible.

The new Section 4.4 describes an algorithm constructed by the author and an electrical engineer, Peyman Milanfar. It employs another algorithm designed for the reconstruction of a convex body from its surface area measure, the topic of the new Section A.4. Two reasons lie behind the choice of this material. Firstly, the first edition was noticeably short of algorithms and devoid of any that apply to the sort of noisy measurements encountered in practice. There is still plenty to be done in this direction. Secondly, it is much easier to tailor one's own work to suit a book than those of others! But if time, energy, and publisher allowed a complete rewriting, there would be a very different book with the same title and author as this one, stating and proving in the text many results of others here only briefly mentioned in chapter notes.

The collaboration with Milanfar followed a web search for the phrase "geometric tomography" made in 1996. I discovered to my surprise that the term had been used independently at least twice after I introduced it at the Oberwolfach meeting on tomography in 1990. The usage in [683, Chapter 7] has essentially no overlap with ours, but that of Thirion [802] is quite close in spirit; he defines geometric tomography to be the process of reconstructing the external or internal boundaries of objects from their X-rays. The same web search led me to the program of Alan Willsky, an electrical engineer at MIT, outlined in Note 1.5.

Much of the work represented by the additional references was presented at various meetings on convex geometry or discrete tomography during the past decade, but several international meetings have featured geometric tomography specifically. Two Summer Schools on Local Stereology and Geometric Tomography were organized by Eva Vedel Jensen at the Sandbjerg Estate, Denmark, the first in 2000 and the second in 2002. In 2004, Salvador Gomis organized the Workshop on Geometric Tomography in Alicante, Spain. It is wonderful to have such energetic and capable friends in beautiful locations.

Terminology and notation are constantly changing in mathematics. Currently "origin symmetric" seems to be favored over "centered," and $\mathbb{R}^n$ is used more and more instead of $\mathbb{E}^n$ as this part of geometry moves into the mainstream. One would think that a notion as basic as volume would enjoy a standard notation, but V, Vol, and vol, with or without subscripts, are all common. In the end I decided to retain most of the terminology and notation of the first edition, so that at least the two editions would be compatible. There are a couple of exceptions: the notation for Radon and Fourier transform has been interchanged, and the definition of Fourier transform is slightly different.

I am obliged to Hugh Murrell for permission to use his Mathematica program that produced Figure 1.11, and to Ulrich Brehm for providing Figure 7.3. To the list of people in the preface to the first edition to whom I owe thanks should be added Robert Huotari, Alex Koldobsky, and Maria Moszyńska. I am especially grateful to Paolo Gronchi, Markus Kiderlen, Wolfgang Weil, and Gaoyong Zhang for their very helpful comments on this edition.

Department of Mathematics
Western Washington University
Bellingham, WA 98225-9063
E-mail: Richard.Gardner@wwu.edu
Web page: http://www.ac.wwu.edu/~gardner

PREFACE

The title of this book, *Geometric Tomography*, is designed to cover the area of mathematics dealing with the retrieval of information about a geometric object from data about its sections, or projections, or both. The term "geometric object" is deliberately vague; a convex polytope or body would certainly qualify, as would a star-shaped body, or even, when appropriate, a compact set or measurable set.

The word "tomography" originates from the Greek $\tau\acute{o}\mu o\varsigma$, meaning a slice. Mathematical computerized tomography is already a recognized subject with an enormously important application in the medical CAT scanner, with which an image of a section of a human patient can be reconstructed from X-rays. Mathematically, the object being reconstructed is a density function, and since it is known that the solution is not unique, no matter how (finitely) many X-rays are taken, the reconstructed picture is always an approximation. When density functions are replaced by geometric objects, there is some hope of a unique solution, and this gives geometric tomography a rather different flavor.

Despite this, there is a point where geometric tomography and computerized tomography merge, and both utilize integral transforms such as the Radon transform. Our definition of geometric tomography is also reminiscent of definitions of stereology or geometric probing. These subjects each have their own distinct viewpoint, while sharing common features with geometric tomography. For example, stereology focuses on random data and statistical methods, but also draws on integral geometry (as do two other related subjects, image analysis and mathematical morphology).

Via these connections and others discussed in the chapter notes, geometric tomography does not lack possible applications.

In computerized tomography, the term "projection" is routinely used for an X-ray. In this book, we adhere to the accepted mathematical definition of a projection as a shadow. In this sense, projections are of little or no interest in computerized tomography. In geometry, however, the well-known polar duality provides a link between sections and projections. There is, in fact, a remarkable correspondence between results concerning projections and those concerning sections through a fixed point, and it would be quite inappropriate to consider the one without considering the other.

Since projections are only shadows, the convex sets are a natural class of objects with which to work. Aristotle's argument, that the earth must be spherical since its shadows on the moon are circular during a lunar eclipse, is in the spirit of geometric tomography. Such matters can often be settled with standard geometrical arguments, but more sophisticated methods are required when only the areas, rather than the exact shapes, of the projections of a convex set are available. Then the Brunn–Minkowski theory, which includes Minkowski's theory of mixed volumes and which forms the core of classical convexity, becomes the ideal framework. In this way, geometric tomography absorbs concepts such as the support function of a convex body, sets of constant width and brightness, zonoids and projection bodies, and projection functions; results such as Aleksandrov's projection theorem and the solution to Shephard's problem; and tools such as Aleksandrov's area measures, the Aleksandrov–Fenchel inequality, and the cosine transform.

Geometric tomography overlaps with convexity, but is not subsumed under it. When the data concern sections through a fixed point, the sets that are star-shaped with respect to that point form a more appropriate class than the convex sets. Within the past three decades or so, a "dual Brunn–Minkowski theory" has arisen, including Erwin Lutwak's dual mixed volumes, and again providing a natural setting. As a consequence, geometric tomography assimilates concepts such as the radial function of a star body, sets of constant section, intersection bodies, and section functions; results such as Funk's section theorem and the solution to the Busemann–Petty problem; and tools such as the i-chord functions, the dual Aleksandrov–Fenchel inequality (a suitable form of Hölder's inequality), and the spherical Radon transform.

The items in the last sentences of the previous two paragraphs are in some sense dual to each other. There is a quite mysterious correspondence in concepts, results, and tools, which polar duality hardly begins to explain.

A parallel X-ray of a body (see Table 1 at the end of this preface) carries more information than a projection. Challenged by P. C. Hammer's 1963 problem, the author and Peter McMullen showed in 1980 that parallel X-rays in certain sets of four directions suffice to determine the shape of any convex body. Later, Aljoa Volčič proved that X-rays emanating from certain sets of four points suffice to determine the shape of any planar convex body. Again, there are two sets of results, for parallel and point X-rays, which are in some sense dual to each other. However, a point X-ray is just a special section function. This provides a bridge to the material described earlier, and dual mixed volumes and the spherical Radon transform are seen in action again.

When the areas of projections, or sections through a point, of a set do not determine it uniquely, one can still hope to estimate its volume. For projections of a convex body, such an estimate is provided by the isoperimetric inequality and Cauchy's surface area formula, but much better estimates are known. For example, Lutwak applied the affinely invariant Petty projection inequality to obtain an estimate in which equality holds for ellipsoids rather than balls. Other affine isoperimetric inequalities furnish corresponding estimates which similarly have the advantage of being invariant under affine transformations.

Geometric tomography houses a zoo of strange geometric bodies, powerful integral transforms, and exotic but highly effective inequalities. Teeming with open problems, it appears an extraordinarily fertile area for research. It resembles particle physics in that symmetry – for example, the duality alluded to before – sometimes allows missing theorems to be predicted, though proofs are not always easy to find. It is too much to expect a Grand Unified Theory, but a more satisfying synthesis is surely within reach.

Some of the open problems listed at the end of each chapter require advanced knowledge and may be very difficult, but others should be quite accessible to undergraduate students. With this in mind, considerable effort has been made to cater to those who have not attended courses in real analysis or convexity. We assume knowledge of calculus, linear algebra, and the basic geometry and topology of two and three dimensions, that is, terms such as scalar product, norm, subspace, interior, boundary, open set, compact set, and connected set. A student with these prerequisites should start with Chapter 1 and, with occasional reference to Chapter 0, be capable of understanding most of it, and nearly all if also familiar with complex numbers and the idea of a metric space. Later chapters involve more advanced topics, but gentle introductions are provided in Chapter 0 and the appendixes. The beginner can make inroads by consulting these and the illustrations, though even without them there should be much that can be absorbed. For example, Chapter 2 follows the same footpath as Chapter 1, with a few brambles created by the appearance of some measure theory. The rest of the book can be read independently of these first two chapters. Chapters 3 and 4 are largely classical convexity; the support function, mixed volumes, and area measures enter here, in such a way that these tools are motivated, rather than required in advance. With Chapters 5 and 6, a new route is followed, needing some measure theory, but mostly set in the plane and again mostly independent of the previous chapters. This route continues in Section 7.2, where dual mixed volumes can be seen at work for the first time. Though much of Chapters 7 and 8 mirrors Chapters 3 and 4, little hangs on the earlier material. Chapter 9, however, draws substantially from both Chapter 4 and Chapter 8.

Many different books could have been written with the title of this one. Here one theme unfolds, supported by the sort of detailed proofs appreciated by most students. Inevitably several important topics, such as Dvoretzky's theorem and others from the local theory of Banach spaces, and the Crofton intersection formulas of integral geometry, are relegated to the chapter notes. Furthermore, Euclidean space and the projective plane hold enough difficulties for us here, though many of the concepts introduced carry over to the more general homogeneous spaces, as in [388], for example.

To add some historical perspective, several biographies have been included in the chapter notes. The history of mathematics is a fascinating subject and a powerful but largely unexploited spur to the learning process. However, accurate historical writing requires special expertise together with careful examination of original documents. The author cheerfully admits his incompetence in this area; the biographies are merely thumbnail sketches pasted together from secondary material.

Though the main text of the book is almost entirely self-contained, the supporting material in Chapter 0 and the appendixes makes frequent reference to the literature. Such references have been limited, when possible, to suitable books (rather than journal articles). There is no escaping the fact that much of the heavy machinery from the Brunn–Minkowski theory eventually comes into play. In the early stages of writing no text contained all the necessary material, and it seemed that the pedestal might be too big for the statue. By great good fortune, Rolf Schneider's comprehensive – and pedagogically sound – treatise [737] on the Brunn–Minkowski theory appeared. The reader who wishes to consult Schneider's volume for more information will find that our notation and terminology are very similar to his.

There are many friends and colleagues to thank. The book evolved from notes of weekly lectures given in late 1989 and early 1990 at the Istituto Matematico "U. Dini" in Florence, during a visit to the Istituto di Analisi Globale e Applicazioni, directed by Professor C. Pucci. This lecture series was confined to Chapter 1 and small parts of Chapters 2 and 5, but was the spark that lit the fire. As work progressed, various assistance was lent by A. D. Aleksandrov, H. Antosiewicz, Keith Ball, John Beem, Yuri Burago, Stefano Campi, Branko Curgus, Hans Debrunner, Hans Goertz, Vladimir Golubyatnikov, Marco Longinetti, Luis Montejano, Frank Morgan, Alain Pajor, Washek Pfeffer, Hans Sagan, Steven Skiena, Alan Thompson, and Tohru Uzawa. Parts of drafts were read by Edoh Amiran, Don Chakerian, Lauren Cowles, Ken Falconer, Paul Goodey, Eric Grinberg, Peter Gritzmann, Helmut Groemer, Peter Gruber, Daniel Hug, Bob Jewett, Hans Kellerer, Joop Kemperman, Dan Klain, Vic Klee, Attila Kuba, Árpád Kurusa, Erwin Lutwak, Horst Martini, Peter McMullen, Frank Natterer, Rolf Schneider, Aljoa Volčič, Wolfgang Weil, and Gaoyong Zhang. Between them they made many valuable suggestions and caught copious mistakes and misprints. (There are infinitely many of these in every manuscript, since each time you look, you find another one.)

Most of the pictures were drawn with Aldus Freehand by Jill Skeels, in a project supported by a grant from the U.S. National Science Foundation. Others

were made by the author, with Mathematica and Virtuoso, and four of these used programs written by Don Chalice (Figure 7.2, the original of which appears in the paper [96] of Ulrich Brehm), Alfred Gray (Figure 3.9, see [330, p. 427]), Branko Grünbaum (Figure 2.1), and Fred Pickel (Figure 8.3, intersection body of the cube). The book was written in LaTeX, part of the wonderful TeX package invented and donated to the world by Donald Knuth.

Several of those already mentioned deserve extra thanks. During a visiting year at the University of California at Davis, Don Chakerian lent his friendly help and encyclopedic knowledge of convexity. At this time the book was to be a joint work with Aljoa Volčič; though circumstances forced him to withdraw from the project, he wrote a first draft of Chapter 5, and our collaboration continued in research incorporated in the text. (This and some other research of the author, mostly also published elsewhere, were partly supported by a grant from the U.S. National Science Foundation.) Rolf Schneider kindly gave me a preprint of his book, which immediately became indispensable. Constant encouragement and expert advice from Erwin Lutwak became a pillar of support as the truth dawned of Gian-Carlo Rota's maxim: When you write a research paper, you are afraid that your results might already be known, but when you write an expository work, you discover that nothing is known.

function	symbol	data	symmetral	symbol
parallel X-ray of K in a direction u	$X_u K$	lengths of chords of K parallel to u	Steiner symmetral	$S_u K$
k-dimensional X-ray of K parallel to a k-dimensional subspace S	$X_S K$	volumes of sections of K parallel to S	k-symmetral	$S_S K$
ith projection function of K	$v_{i,K}$	volumes of projections of K on each i-dimensional subspace	exists only for $i = 1$ or $(n-1)$	
1st projection function, or width function, of K	$w_K = v_{1,K}$	lengths of projections of K on each line through the origin	central symmetral	ΔK
$(n-1)$th projection function, or brightness function, of K	$v_K = v_{n-1,K}$	volumes of projections of K on each hyperplane through the origin	Blaschke body	∇K
ith section function of K, or i-dimensional X-ray of K at the origin	$\tilde{v}_{i,K}$	volumes of sections of K by each i-dimensional subspace	i-chordal symmetral	$\tilde{\nabla}_i K$
1st section function of K, or X-ray of K at the origin	$X_o K = \tilde{v}_{1,K}$	lengths of sections of K by lines through the origin	chordal symmetral	$\tilde{\Delta} K = \tilde{\nabla}_1 K$
section function, or $(n-1)$-dimensional X-ray at the origin	$\tilde{v}_K = \tilde{v}_{n-1,K}$	volumes of sections of K by each hyperplane through the origin	dual Blaschke body	$\tilde{\nabla} K = \tilde{\nabla}_{n-1} K$
X-ray of K at a point p	$X_p K = X_o(K-p)$	lengths of sections of K by lines through p	chordal symmetral at p	$\tilde{\Delta}_p K = \tilde{\Delta}(K-p)$

Note: In the table, K is a convex body; however, most of the functions and symmetrals are defined for more general sets. Each symmetral is a symmetric body yielding the same data from the corresponding function as K.

Table 1. Some functions and symmetrals of geometric tomography.

0

Background material

This chapter introduces notation and terminology and summarizes aspects of the theories of affine and projective transformations, convex and star sets, and measure and integration appearing frequently in the sequel.

Some passages are designed to ease the beginner into these areas, but not all the material is elementary. It is intended that the reader **start with Chapter 1**, and use the present chapter as a reference manual. For Chapter 1, *the requisite material is included in the first four sections of this chapter only*, and for Chapter 2, *the requisite material is included in the first five sections only*.

0.1. Basic concepts and terminology

This section is a brief review of some basic definitions and notation. Any unexplained notation can be found in the list at the end of the book.

Almost all the results in this book concern Euclidean n-dimensional space $\mathbb{E}^n$. The origin in $\mathbb{E}^n$ is denoted by o, and if $x \in \mathbb{E}^n$, we usually label its coordinates by $x = (x_1, \ldots, x_n)$. (In $\mathbb{E}^2$ and $\mathbb{E}^3$ we often use a different letter for a point and label its coordinates in the traditional way by x, y, and z.) The Euclidean norm of x is denoted by $\|x\|$, and the Euclidean scalar product of x and y by $x \cdot y$. The closed line segment joining x and y is $[x, y]$. Points are identified with vectors, and are always denoted by lowercase letters. For sets we usually employ capitals, although we also use lowercase for straight lines. Script capitals are used for classes of sets; an exception is the $\mathcal{S}$ we use for sets of directions in Chapters 1 and 2, but here we are really identifying a direction with the line through the origin parallel to it. The natural numbers, real numbers, and complex numbers have the usual symbols $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$. The letters i, j, k, m, and n denote integers unless it is stated otherwise (in parts of the book i often represents a real number), or unless we are working with complex numbers, when $i^2 = -1$

as usual. In particular, the default meaning of an expression such as $1 \le i \le n$ is $i \in \{1, \dots, n\}$.

The *unit ball* in $\mathbb{E}^n$ is $B = \{x : \|x\| \le 1\}$, with surface the *unit n-sphere* $S^{n-1} = \{x : \|x\| = 1\}$. When necessary we may write B^n instead of B. We attempt to reserve u for the members of S^{n-1}, the unit vectors. If $u \in S^{n-1}$, then $u^\perp$ is the $(n-1)$-dimensional subspace orthogonal to u, and l_u the 1-dimensional subspace parallel to u. Generally, S is used for a subspace, and $S^\perp$ for its complementary orthogonal subspace. The Grassmann manifold of k-dimensional subspaces of $\mathbb{E}^n$ is denoted by $\mathcal{G}(n, k)$. More often than not the topology on $\mathcal{G}(n, k)$ is unnecessary, and the symbol then simply denotes the corresponding set of subspaces.

Translates of subspaces are called *planes* or *flats*, or *hyperplanes* if they are $(n-1)$-dimensional. A hyperplane divides the space into two *half-spaces* (*half-planes* in $\mathbb{E}^2$). A *ray* is a semi-infinite straight line. If E is a set, the *linear hull* $\operatorname{lin} E$ and *affine hull* $\operatorname{aff} E$ of E are, respectively, the smallest subspace and the smallest plane containing E. The *dimension* $\dim E$ of a set E is the dimension of its affine hull.

We say that two planes are *parallel* if one is contained in a translate of the other, and *orthogonal* if, when translated so that they contain the origin, one contains the complementary orthogonal subspace of the other. (These terms are often used by other authors in a more restrictive way.) A *slab* is the closed region between two parallel hyperplanes.

Suppose that F_1, F_2 are planes in $\mathbb{E}^n$, of dimensions d_1 and d_2, respectively. Then by [85, Theorem 32.1], either $F_1 \cap F_2 = \emptyset$ or $\dim(F_1 \cap F_2) \ge d_1 + d_2 - n$. The planes F_1 and F_2 are in *general position* with respect to each other if either $d_1 + d_2 < n$, $F_1 \cap F_2 = \emptyset$, and there is no direction parallel to both planes, or $d_1 + d_2 \ge n$ and $\dim(F_1 \cap F_2) = d_1 + d_2 - n$. See [85, pp. 88–90] for more information. A finite set of points in $\mathbb{E}^n$ is said to be in *general position* if no more than $k + 1$ of them belong to any k-dimensional plane.

A few of our results are set in 2-dimensional projective space $\mathbb{P}^2$. Generally, *n-dimensional projective space* $\mathbb{P}^n$ can be defined as the space of 1-dimensional subspaces of $\mathbb{E}^{n+1}$. The points of $\mathbb{P}^n$ are labeled by *homogeneous coordinates* $w = (w_1, \dots, w_{n+1})$, not all zero, so for real $t \ne 0$ the points w and tw are identified; see, for example, [85, p. 217]. In this way, $\mathbb{P}^1$ can be regarded as the unit circle S^1 with antipodal points identified. We can also identify $\mathbb{E}^n$ with $\{w : w_{n+1} \ne 0\}$, where the usual coordinates are given by $x_i = w_i / w_{n+1}$. The remaining set $H_\infty = \{w : w_{n+1} = 0\}$ is the *hyperplane at infinity* (strictly speaking, a copy of $\mathbb{P}^{n-1}$). In particular, $\mathbb{P}^2$ can be regarded as $\mathbb{E}^2$ with a *line at infinity* (strictly speaking, a copy of $\mathbb{P}^1$) adjoined.

Our terminology for set theory and topology is standard. If E is a set, then $|E|$, $\operatorname{co} E$, $\operatorname{cl} E$, $\operatorname{int} E$, and $\operatorname{bd} E$ denote the *cardinality*, *complement*, *closure*, *interior*, and *boundary* of E, respectively; also, $\operatorname{relint} E$ is the *relative interior* of E, that is, the interior of E relative to aff E. The *relative boundary* of E is the boundary of E relative to aff E. The *symmetric difference* of E and F is

$$E \triangle F = (E \setminus F) \cup (F \setminus E).$$

A G_δ *set* is a countable intersection of open sets, and an F_σ *set* is a countable union of closed sets. A set is of *first category* if it is the countable union of nowhere dense sets, and *residual* if it is the complement of a set of first category. A set in a locally compact Hausdorff space is residual if it contains a dense G_δ set; see, for example, [700, pp. 158–60 and 200–1]. A *component* of a set is a maximal connected subset. A closed set is *regular* if it is the closure of its interior, and a *body* is a compact, regular set.

The *diameter* diam E of a set E is

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

If x is a point and E is a closed set, the *distance* between x and E is

$$d(x, E) = \inf\{\|x - y\| : y \in E\}.$$

If E and F are sets, and r is a real number, then

$$E + F = \{x + y : x \in E, y \in F\},$$

and

$$rE = \{rx : x \in E\}.$$

A set E is called *centered* if $-x \in E$ whenever $x \in E$, and *centrally symmetric* if there is a vector c such that the translate $E - c$ of E by $-c$ is centered. In the latter case c is called a *center* of E. The center of a nonempty bounded centrally symmetric set is unique.

If X is a subset of $\mathbb{E}^n$, or indeed any topological space, the *support* of a real-valued function f on X is the set $\operatorname{cl}\{x \in X : f(x) \neq 0\}$. We denote by $C(X)$ the class of continuous real-valued functions on X. When X is an appropriate subset of $\mathbb{E}^n$, $C_e(X)$ denotes the even functions in $C(X)$, and $C_e^+(X)$ the nonnegative functions in $C_e(X)$.

If f and g are real-valued functions, we say that $f = O(g)$ on $A \subset \mathbb{R}$ if there is a constant c such that $|f(x)| \leq c|g(x)|$ for all $x \in A$. When $A = \mathbb{N}$, we sometimes say that $f = O(g)$ as $n \to \infty$, while $f = O(g)$ as $x \to 0$ means that $f = O(g)$ on $A = (0, a)$ for sufficiently small a.

0.2. Transformations

No single book seems to provide a completely satisfactory introduction to the various types of transformations of $\mathbb{E}^n$ and $\mathbb{P}^n$; somehow the required material falls between the texts on Euclidean or projective geometry currently available. Borsuk's book [85] is possibly the most comprehensive text for this purpose, but its notation is quite outdated.

If A is an $n \times n$ matrix, the inverse and transpose of A are denoted by A^{-1} and A^t. We call A *singular* or *nonsingular* according to whether $\det A = 0$ or $\det A \neq 0$, respectively; A^{-1} exists precisely when A is nonsingular. We also adopt the abbreviation A^{-t} for $(A^{-1})^t$. Note that if A is nonsingular, then A^t is also, and $(A^t)^{-1} = (A^{-1})^t$.

For transformations ϕ of $\mathbb{E}^n$ and $\mathbb{P}^n$, we shall permit ourselves the shorthand $\phi x = \phi(x)$. The reader may find Figure 0.1 useful in interpreting the definitions given below.

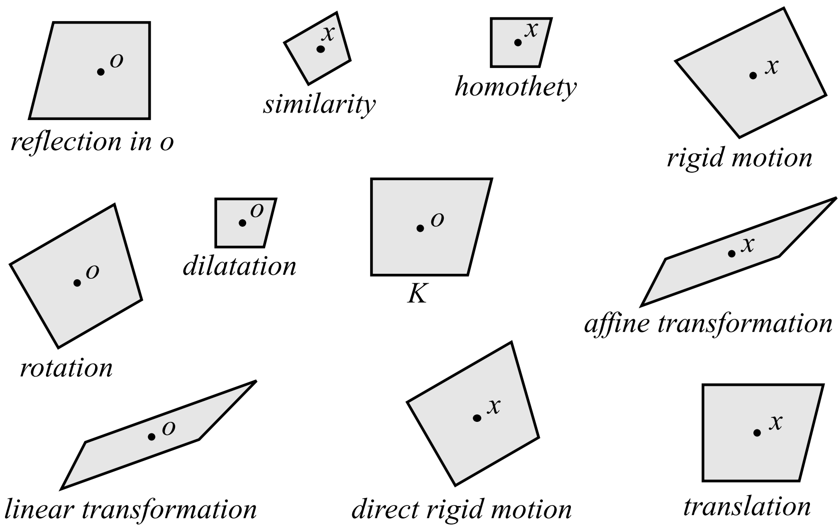


Figure 0.1. Transformations of a set K.

A *linear transformation* (or *affine transformation*) of $\mathbb{E}^n$ is a map ϕ from $\mathbb{E}^n$ to itself such that $\phi x = Ax$ (or $\phi x = Ax + t$, respectively), where A is an $n \times n$ matrix and $t \in \mathbb{E}^n$. (Here x is considered as a column vector, of course.) We call ϕ *singular* or *nonsingular* according to whether A is singular or nonsingular, respectively. The group of nonsingular linear (or affine) transformations is denoted by GL_n (or GA_n); its members are, in particular, bijections of $\mathbb{E}^n$ onto itself. The group of *special linear* (or *special affine*) transformations of $\mathbb{E}^n$ is denoted by SL_n (or SA_n, respectively). These are the members of GL_n (or GA_n) whose determinant is one. We shall write $\det \phi$ instead of $\det A$, and ϕ^{-1}, ϕ^t, and ϕ^{-t} for the affine transformations with corresponding matrices A^{-1}, A^t, and A^{-t}, respectively.

If A is the identity matrix, then $\phi x = x + t$, and the map ϕ is called a *translation*. Each affine transformation is composed of a linear transformation followed by a translation.

Any set of $n + 1$ points in general position in $\mathbb{E}^n$ can be mapped onto any second set of $n + 1$ points by a suitable affine transformation, and the latter is nonsingular if the second set is also in general position (see [595, Theorem 7, p. 16]).

If $\phi \in GA_n$, then ϕ takes parallel k-dimensional planes onto parallel k-dimensional planes (cf. [85, p. 156]).

An *isometry* of $\mathbb{E}^n$ is a map ϕ such that $\|\phi x - \phi y\| = \|x - y\|$; in other words, a distance-preserving bijection. Isometries are also called congruences, and the image and pre-image under an isometry are said to be *congruent*. Every isometry is affine (see, for example, [85, p. 150] or [839, p. 139]). Examples of isometries are the translations and the *reflections*, which map all points to their mirror images in some fixed point, line, or plane. (In particular, $\phi x = -x$ is the reflection in the origin.)

If $F = S + x_0$ (where $S \in \mathcal{G}(n, k)$, $x_0 \in \mathbb{E}^n$, and $1 \leq k \leq n - 1$) is a k-dimensional plane, and $x \in \mathbb{E}^n$, then there are unique points $y \in S$ and $z \in S^\perp$ such that $x = y + z$, and we can define a map taking x to $y + x_0 \in F$. This map is the (orthogonal) *projection* on the plane F. It is a singular affine transformation. If E is an arbitrary subset of $\mathbb{E}^n$, the image of E under a projection on a plane F is called the *projection of E on F* and denoted by $E|F$. Since $E|S$ is a translate of $E|F$ when $F = S + x_0$, we almost always work with the former.

If $\phi \in GL_n$, then

$$x \cdot \phi y = \phi^t x \cdot y, \tag{0.1}$$

for all $x, y \in \mathbb{E}^n$. The *orthogonal group O_n* of orthogonal transformations consists of those isometries of $\mathbb{E}^n$ that are also linear transformations; these are precisely the maps ϕ preserving the scalar product, that is, $\phi x \cdot \phi y = x \cdot y$. (An orthogonal matrix satisfies $A^t = A^{-1}$ and by (0.1) we have $\phi^t = \phi^{-1}$, hence the name.) It follows from this that orthogonal transformations have determinants with absolute value one. As is shown in [85, Theorem 50.6], every isometry is an orthogonal transformation followed by a translation, and for this reason isometries are sometimes also called *rigid motions*. The *special orthogonal group SO_n* of *rotations* about the origin consists of those orthogonal transformations with determinant one. A *direct rigid motion* is a rotation followed by a translation; these do not allow reflection.

A *dilatation* is a map $\phi x = rx$, for some $r > 0$. A *homothety* is a map $\phi x = rx + t$, for some $r > 0$ and $t \in \mathbb{E}^n$, that is, a composition of a dilatation with a translation (this is sometimes referred to as a direct homothety). A *similarity* is a composition of a dilatation with a rigid motion. We say two sets are *homothetic* (or *similar*) if one of them is an image of the other under a homothety (or similarity, respectively), or if one of the sets is a single point.

We find occasional use for projective transformations of $\mathbb{P}^n$. Such a transformation is given in terms of homogeneous coordinates by $\phi w = Aw + t$, where A is an $(n+1) \times (n+1)$ matrix and $t \in \mathbb{E}^{n+1}$, and where ϕ is called nonsingular if $\det A \neq 0$. Since we can regard $\mathbb{P}^n$ as $\mathbb{E}^n$ with a hyperplane H_∞ adjoined, we can also speak of a projective transformation of $\mathbb{E}^n$. In this regard, another formulation is useful. A *projective transformation* ϕ of $\mathbb{E}^n$ has the form

$$\phi x = \frac{\psi x}{x \cdot y + t}, \tag{0.2}$$

where $\psi \in GA_n$, $y \in \mathbb{E}^n$, and $t \in \mathbb{R}$, and ϕ is nonsingular if the associated linear map

$$\bar{\psi}(x, 1) = (\psi x, x \cdot y + t)$$

is nonsingular. If $y = o$, then ϕ is affine, but if $y \neq o$, ϕ maps the hyperplane $H = \{x : x \cdot y + t = 0\}$ onto H_∞. To avoid points in a set E being mapped into H_∞, we may insist that ϕ be *permissible* for E; this simply means that $E \cap H = \emptyset$.

Projective transformations map planes onto planes (neglecting the points mapping to or from infinity); see [595, pp. 19–20]. They also preserve cross ratio; a proof is given in [85, Corollary 96.11]. (The *cross ratio* of four points x_i, $1 \leq i \leq 4$ on a line is defined by

$$\langle x_1, \ldots, x_4 \rangle = \frac{(x_3 - x_1)(x_4 - x_2)}{(x_4 - x_1)(x_3 - x_2)},$$

where x_i also denotes the coordinate of the point x_i in a fixed Cartesian coordinate system on the line.) Affine transformations are also projective transformations, so the former also preserve cross ratio.

The sets E and F are called *linearly, affinely*, or *projectively equivalent* if there is a nonsingular transformation ϕ, linear, affine, or projective and permissible for E, respectively, such that $\phi E = F$. Suppose that E and F are bounded centered sets affinely equivalent via a nonsingular transformation ϕ. If $\phi o = p$, then p is the center of F; but since o is the unique center of F, we have $p = o$. Therefore ϕ is linear, proving that E and F are linearly equivalent.

0.3. Basic convexity

There are several possibilities for an introduction to the basic properties of convex sets. For the absolute beginner, the books of Lay [499] and Webster [827] are recommended. The first chapter of [595], by McMullen and Shephard, is terse, but very informative, as is the first chapter of [737], by Schneider. The text of [845], by Yaglom and Boltyanskiĭ, is set out in the form of exercises and solutions, with plenty of helpful diagrams. Chapters 11 and 12 of Berger's two-volume set [52], [53], contain some wonderful pictures, and Lyusternik's little book [554] is quirky but delightful. A list of books on convexity can be found in [737, p. 433].

A set C in $\mathbb{E}^n$ is called *convex* if it contains the closed line segment joining any two of its points, or, equivalently, if $(1-t)x + ty \in C$ whenever $x, y \in C$ and $0 \le t \le 1$. A convex set, then, has no "holes" or "dents." A *convex body* is a compact convex set whose interior is nonempty; this definition conforms with general usage, but the reader is warned that in the important texts of Bonnesen and Fenchel [83] and Schneider [737] any compact convex set qualifies as a convex body. The *convex hull* $\operatorname{conv} E$ of a set E is the smallest convex set containing it.

If C is a compact convex set, a *diameter* of C is a chord $[x, y]$ of C such that $\|x - y\| = \operatorname{diam} C$.

A hyperplane H *supports* a set E at a point x if $x \in E \cap H$ and E is contained in one of the two closed half-spaces bounded by H. We say H is a *supporting hyperplane* of E if H supports E at some point.

A convex body is *strictly convex* if its boundary does not contain a line segment and *smooth* if there is a unique supporting hyperplane at each point of its boundary.

The intersection of a compact convex set with one of its supporting hyperplanes is called a *face*, and $(n-1)$-dimensional faces are also called *facets*. An *extreme point* of K is one not contained in the relative interior of any line segment contained in K. The point x is called an *exposed point* of K if there is a supporting hyperplane H such that $H \cap K = \{x\}$. Every exposed point is extreme, but the converse is not true. Also, a compact convex set is the closure of the convex hull of its exposed points, implying that every compact convex set has at least one exposed point (see [737, Section 1.4], especially Theorem 1.4.7). A *corner point* of a compact convex set in $\mathbb{E}^2$ is one at which there is more than one supporting line.

If K_1 and K_2 are disjoint compact convex sets in $\mathbb{E}^n$, then there is a hyperplane H that (strictly) *separates* K_1 and K_2; that is, K_1 is contained in one open half-space bounded by H, and K_2 in the other. A proof can be found in [499, Theorem 4.12] or [737, Theorem 1.3.7]. (In infinite-dimensional spaces, this separation theorem is closely related to the Hahn–Banach theorem; see [52, Section 11.4].)

Every affine transformation preserves convexity. If ϕ is a projective transformation, permissible for a line segment, then it maps this line segment onto another line segment. Therefore ϕ preserves the convexity of convex bodies for which it is permissible.

A nonempty subset C of $\mathbb{E}^n$ is a *cone* with vertex o if $ty \in C$ whenever $y \in C$ and $t \ge 0$. A *convex cone* with vertex o is a cone with vertex o that is convex; such a set is closed under nonnegative linear combinations. A cone (or convex cone) with vertex x is of the form $C + x$, where C is a cone (or convex cone, respectively) with vertex o.

Let us define some special convex bodies. The unit ball B in $\mathbb{E}^n$ was defined already. A *ball* is any set homothetic to B, and an *ellipsoid* is an affine image of B. The centered n-dimensional ellipsoids whose axes are parallel to the coordinate axes are of the form

$$\left\{x : \sum_{i=1}^{n} \frac{x_i^2}{a_i^2} \leq 1\right\}.$$

If $0 \leq k \leq n$, a *k-dimensional simplex* in $\mathbb{E}^n$ is the convex hull of $k+1$ points in general position.

A *polyhedron* is a finite union of simplices; in $\mathbb{E}^2$, we shall use the term *polygon* instead. A convex polyhedron or *convex polytope* can also be defined as the convex hull of a finite set of points. We denote by $\mathcal{F}_k(P)$ the set of k-dimensional faces of a convex polytope P.

Important examples of convex polytopes are the *unit cube* $\{x : 0 \leq x_i \leq 1, 1 \leq i \leq n\}$ (and *centered unit cube* $\{x : |x_i| \leq 1/2, 1 \leq i \leq n\}$) in $\mathbb{E}^n$; the *parallelepipeds* or *parallelotopes*, affine images of the unit cube; the *boxes*, rectangular parallelepipeds with facets parallel to the coordinate hyperplanes; and the *cross-polytopes* (n-dimensional versions of the octahedron), each the convex hull of n mutually orthogonal line segments sharing the same midpoint. An n-dimensional *pyramid* P is the convex hull of an $(n-1)$-dimensional convex polytope Q (its *base*) and a point $x \notin \operatorname{aff} Q$ called the *apex* of P.

A (right spherical) *cylinder* in $\mathbb{E}^n$ is the Cartesian product of an $(n-1)$-dimensional ball C and a line segment orthogonal to $\operatorname{aff} C$. A (right spherical) *bounded cone* in $\mathbb{E}^n$ is the convex hull of an $(n-1)$-dimensional ball C and a point on the line orthogonal to $\operatorname{aff} C$ through the center of C.

If K is a convex body in $\mathbb{E}^n$, we denote by $r(K)$ and $R(K)$ the *inradius* and *circumradius* of K. These are the radii of the largest n-dimensional ball contained in K and the smallest ball containing K, respectively.

Topologically, a convex body is not very interesting. The surface of a convex body K in $\mathbb{E}^n$ is homeomorphic to S^{n-1} via a *radial map* f, defined by selecting a point $x_0 \in \operatorname{int} K$ and letting

$$f(x) = (x - x_0)/\|x - x_0\|, \tag{0.3}$$

for each $x \in \operatorname{bd} K$.

A real-valued function on $\mathbb{E}^n$ is *convex* if

$$f\big((1-t)x + ty\big) \leq (1-t)f(x) + tf(y),$$

for all $x, y \in \mathbb{E}^n$ and $0 \leq t \leq 1$, and *concave* if $-f$ is convex. (The terms concave up and concave down are sometimes used for convex and concave, respectively.)

0.4. The Hausdorff metric

Exactly what does it mean to say that a sequence of compact sets converges to another compact set? One must have a way of measuring the distance between two compact sets. This notion of distance must behave like the usual distance $d(x, y) = \|x - y\|$ between points, which has three fundamental properties: $d(x, y) \geq 0$, and equals zero if and only if $x = y$; $d(x, y) = d(y, x)$; and the triangle inequality

$$d(x, z) \leq d(x, y) + d(y, z).$$

Such a function is called a *metric*. We shall only define one metric for compact sets here, though there are several in common use (see Lemma 1.2.14 for another). The *Hausdorff metric* δ on the class of nonempty compact sets in $\mathbb{E}^n$ is defined by

$$\delta(E, F) = \max\{\max_{x \in E} d(x, F),\ \max_{x \in F} d(x, E)\}. \tag{0.4}$$

(A geometrically more appealing definition is given later.) It can be checked that δ satisfies the three conditions listed earlier. The proof, and basic properties of the metric space of compact sets in $\mathbb{E}^n$ defined in this way, may be found in [499, Section 14] or [737, Section 1.8].

Suppose that E is a nonempty set in $\mathbb{E}^n$ and $\varepsilon > 0$. Then

$$E_\varepsilon = E + \varepsilon B = \cup_{x \in E} (x + \varepsilon B) \tag{0.5}$$

is called an *outer parallel set* of E. When E is closed, E_ε is just the set of all points whose distance from E is no more than ε. (See [499, Section 14], [737, p. 134]; see also the illustration in the book [789, Fig. 1.1(b)] of Stoyan, Kendall, and Mecke, and the interesting accompanying discussion on the utility of this idea in the processing of images.) This convenient concept allows the following alternative definition of the Hausdorff metric:

$$\delta(E, F) = \min\{\varepsilon > 0 :\ E \subset F_\varepsilon \text{ and } F \subset E_\varepsilon\}. \tag{0.6}$$

This means that the Hausdorff distance between two convex bodies K_1 and K_2 is at most ε if K_1 is contained in the outer parallel body $K_2 + \varepsilon B$ of K_2, and K_2 is contained in the outer parallel body $K_1 + \varepsilon B$ of K_1.

The Hausdorff metric is the standard one in the study of convex sets. We denote by $\mathcal{K}^n$ (or $\mathcal{K}^n_0$) the space of nonempty compact convex sets (or convex bodies, respectively) in $\mathbb{E}^n$ with the Hausdorff metric. (The definition of a body in Section 0.1 implies the existence of interior points when the set is nonempty.) It is the default metric, always used unless stated otherwise, for example, when discussing continuity of a function defined on the class of compact convex sets. A specific, and important, example of this is the continuity of volume on $\mathcal{K}^n$; see [499, Theorem 22.6] or [737, Theorem 1.8.16]. (One should try not to be blasé about such statements. After all, length is not continuous in $\mathbb{E}^2$, since one can approximate a closed line segment arbitrarily closely by polygonal arcs whose lengths are unbounded. According to Young [853, p. 303], this disturbed Lebesgue greatly when he was at school! In fact, length is only semicontinuous in $\mathbb{E}^2$.)

A very frequently quoted theorem is the following one, whose proof may be found in [499, Section 15] or [737, Theorem 1.8.6].

Theorem 0.4.1 (Blaschke's selection theorem). *Every bounded sequence of compact convex sets has a subsequence converging to a compact convex set.*

(A sequence of sets is *bounded* if there is a ball containing each member of the sequence.) In [737, Theorems 1.8.13 and 1.8.15], it is shown that each $K \in \mathcal{K}^n$ can be approximated arbitrarily closely from within or without by convex polytopes. This implies that the class of convex polytopes is dense in $\mathcal{K}^n$. It is also known that both the class of smooth convex bodies and the class of strictly convex bodies are dense in $\mathcal{K}^n$; see [737, Theorem 2.6.1].

0.5. Measure and integration

Measure theory deals with the definition and generalizations of the intuitive notions of length, area, and volume. The subject is amply supplied with well-written books appropriate for the novice. Many a student has learned the basics of Lebesgue measure and integration and the rudiments of general measure theory from [700], by Royden. At a slightly higher level, Munroe's book [639] is to be recommended. Unfortunately, however, the *geometric* aspects of measure theory are often ignored in the standard introductory texts. Exceptions are [839], by Weir (see Chapter 6 of Volume 1), and [410], by Jones (see Chapter 3). Of course, there are books on geometric measure theory proper, but here we can only suggest a browse of the first three of chapters of the entertaining and exquisitely illustrated introduction [637] by Morgan; we use no advanced geometric measure theory in this book.

In practice one can get by without most of the complicated theory of abstract measure. We summarize here the ingredients used in the sequel.

Consider, as a first example, area in the plane. Its essential properties are:

1. Familiar sets such as triangles, disks, and so on can be assigned a real number representing the area of the set.

2. The area of a countable union of disjoint sets is the sum of the areas of the sets; that is, area is *countably additive*.

3. The area of a set does not change when it is moved by a translation; that is, area is *translation invariant*. In fact, area is even invariant under isometries.

The same properties hold for a generalized notion of length in the real line, or volume in space. Length and area are denoted by λ_1 and λ_2, respectively. For Chapter 1, this is all one really needs.

Sooner or later, it becomes necessary to talk about the area of less familiar sets. It turns out that in order to retain the second and third properties, one has to give up the hope of assigning an area to *all* subsets of the plane (at least, if one wishes to use the commonly accepted axiom of choice). However, it can be shown that the concept of area can be defined so that all open sets can be assigned an area. Moreover, one can prove that the family of all sets that can be assigned an area forms a *σ-algebra*; that is, the family contains the empty set and is closed under the taking of complements and countable unions (and therefore also differences and countable intersections). Since the family of *Borel sets* is, by definition, the smallest σ-algebra containing the open sets, all Borel sets can be assigned an area.

Again, the same comments apply to generalized length in the real line and volume in space. Generalized length, area, and volume are examples of measures, and the sets that can be assigned a generalized length, area, or volume are called measurable sets. Among the measurable sets are those *of measure zero*, including all countable sets, but also many uncountable sets. For example, the Cantor ternary set in the real line has zero generalized length, and any line segment in the plane has zero area. Sets of measure zero (sometimes called *null sets*) are often neglected in measure theory, just as the number zero can be ignored in addition. For the types of measures encountered in this book, one is never too far from sanity when working with measurable sets, for it can be shown that each measurable set is the union of countably many closed sets and a (necessarily measurable) set of measure zero.

We are now ready for the formal definitions which abstract these ideas.

Let X be a set. A countably additive, extended real-valued function defined on a σ-algebra of subsets of X is called a *signed measure*; it is a *measure* if it is also nonnegative. The members of the σ-algebra are called *measurable sets*. We say a measure μ is *σ-finite* if X is a countable union of sets of finite μ-measure. A measure μ is said to be *concentrated* on a subset E of X if $\mu(X \setminus E) = 0$. If X is a topological space, and the σ-algebra consists of the Borel sets in X, the measure is called a *Borel measure*. An arbitrary measure in X is called *Borel regular* if Borel sets are measurable and every measurable set is contained in a Borel set of the same measure. A property is said to hold *μ-almost everywhere* or for *μ-almost all $x \in X$* if there is a subset E of X with $\mu(E) = 0$ such that the property holds for all $x \in X \setminus E$.

We generally use lowercase Greek letters for measures. This is the convention adopted by most measure theorists, with the important exception of some who work in geometric measure theory, who use capital script letters, such as the $\mathcal{H}$ for Hausdorff measure (to be defined shortly). History has forced us to make, reluctantly, an exception for the area measures, defined in Section A.2.

After measures are defined, one can deal with the integral (some authors reverse this process). If μ is a measure in X, the *μ-measurable* extended real-valued functions are those for which the inverse image of an open set is a measurable set.

When X is a topological space, there is also the class of *Borel functions* on X, the extended real-valued functions for which the inverse image of an open set is a Borel set. Every continuous function is Borel, and if μ is a Borel measure, then every Borel function is μ-measurable. For certain functions f on X, a meaning can be given to

$$\int_E f(x)\, d\mu(x),$$

the *integral* of f over the measurable set $E \subset X$, in such a way that in the familiar case of a nonnegative f defined on $\mathbb{E}^n$, the integral gives the volume under the graph of f. Nonnegative functions are called *μ-integrable* on E if they are μ-measurable and the integral exists and is finite. An arbitrary function f is *μ-integrable* if both its *positive part* f^+ and its *negative part* f^-, defined by

$$f^+(x) = \max\{f(x), 0\} \text{ and } f^-(x) = \max\{-f(x), 0\},$$

are integrable. A bounded measurable function is integrable on any set of finite measure. All this can be found in Chapters 4 and 11 of [700], for example.

One theorem in the theory of integration is of outstanding importance: Fubini's theorem (see [700, Theorem 19, p. 307]) says that in all reasonable circumstances, the integral of a function on a product of two spaces equals both of the two iterated integrals. (This allows, for example, the volume of a measurable set in $\mathbb{E}^3$ to be calculated by integrating the areas of its sections by planes parallel to a given plane.)

The *n-dimensional Lebesgue measure* λ_n in $\mathbb{E}^n$ is often defined to be the unique Borel-regular, translation-invariant measure in $\mathbb{E}^n$ such that the unit cube has unit measure. This provides one definition of generalized length in the real line, area in the plane, and volume in space. Defined this way, however, λ_n is not the most important measure. This honor goes to *k-dimensional Hausdorff measure* $\mathcal{H}^k$ in $\mathbb{E}^n$, $0 \leq k \leq n$. This is the standard way of measuring k-dimensional volume in $\mathbb{E}^n$, so that, for example, one could use $\mathcal{H}^1$ to measure the perimeter of a disc, or $\mathcal{H}^2$ for the surface area of a ball. The definition of Hausdorff measure (see the texts of Morgan [637, p. 8] or Rogers [694, Chapter 2]) is somewhat technical, but not really more so than the very commonly adopted definition of Lebesgue measure in the real line via Lebesgue outer measure, as in Chapter 3 of [700], for example.

It is a convenient fact that the two measures λ_n and $\mathcal{H}^n$ agree in $\mathbb{E}^n$ (see [637, Corollary 2.8] or [694, Theorem 30]), provided the correct constant is included in the definition of $\mathcal{H}^n$. There is a similar agreement between $\mathcal{H}^{n-1}$ and *$(n-1)$-dimensional spherical Lebesgue measure* in S^{n-1}, the unique Borel-regular, rotation-invariant measure in $\mathbb{E}^n$ such that S^{n-1} has measure equal to the constant ω_n whose value is given by (0.10). Indeed, it is well known that $\mathcal{H}^{n-1}$ is Borel regular and rotation invariant (see [694, Theorem 27 and p. 58]), and the fact that $\mathcal{H}^{n-1}(S^{n-1}) = \omega_n$ follows from integration via the area formula in [637, 3.7, p. 25]. Therefore we allow ourselves to speak loosely of k-dimensional Lebesgue measure in $\mathbb{E}^n$ when we really mean k-dimensional Hausdorff measure, and use λ_k for integration in planes or spheres. Two abbreviations should be noted: We shall write dx (or du, etc.) for $d\lambda_k(x)$ (or $d\lambda_k(u)$, etc., as appropriate) when integrating over a k-dimensional plane or unit $(k+1)$-sphere S^k.

The measure $\mathcal{H}^0$ (we shall write λ_0) is just the counting measure, which counts the number of points in a set.

When no misunderstanding can arise – for example, when working with compact convex sets – we call the λ_k-measure of a k-dimensional body in $\mathbb{E}^n$ its *volume*. This is traditional in geometry.

Often we want to work with the equivalence classes of measurable sets modulo sets of measure zero, and here it is useful to write $E \simeq F$ when $\lambda_n(E \triangle F) = 0$.

Let $\phi \in GA_n$. Then $|\det \phi|$ is the factor by which ϕ changes volume, that is,

$$\lambda_n(\phi E) = |\det \phi| \lambda_n(E), \tag{0.7}$$

for each λ_n-measurable set E in $\mathbb{E}^n$; see [839, pp. 142–4]. It follows that the members of SA_n, and more generally those maps in GA_n whose determinants are ± 1, are volume preserving. It also follows that if $r \geq 0$, then $\lambda_n(rE) = r^n \lambda_n(E)$. More generally, if $1 \leq k \leq n$, E is a λ_k-measurable set in $\mathbb{E}^n$, and $r \geq 0$, then

$$\lambda_k(rE) = r^k \lambda_k(E). \tag{0.8}$$

One can also check that ϕ preserves the ratio of λ_k-measures of sets in parallel k-dimensional planes.

The volume of the unit ball in $\mathbb{E}^n$ is given by

$$\kappa_n = \lambda_n(B) = \frac{\pi^{n/2}}{\Gamma(1 + \frac{n}{2})}, \tag{0.9}$$

with the convention $\kappa_0 = 1$, and its surface area is

$$\omega_n = \lambda_{n-1}(S^{n-1}) = n\kappa_n. \tag{0.10}$$

The first computation is given in [570, pp. 324–5] and the second in [171, p. 125]; or see [746, p. 18]. To calculate special values of κ_n, one only needs $\Gamma(1+x) = x\Gamma(x)$, $\Gamma(1) = 1$, and $\Gamma(1/2) = \sqrt{\pi}$. It is interesting that κ_n increases with n to its maximum value $8\pi^2/15$ when $n = 5$, and then decreases, approaching zero.

Using (0.9) and (0.7), one shows that the n-dimensional centered ellipsoid $\{x : \sum_{i=1}^n x_i^2/a_i^2 \leq 1\}$ has volume

$$a_1 a_2 \cdots a_n \kappa_n. \tag{0.11}$$

The volume of a parallelepiped is the λ_{n-1}-measure of its base times its height orthogonal to its base. The volume of the parallelepiped in $\mathbb{E}^n$ with vertices at o, $p_1, \ldots, p_n$ is also given by

$$|\det(p_{ij})|, \tag{0.12}$$

where $p_i = (p_{i1}, \ldots, p_{in})$, and the volume of the simplex in $\mathbb{E}^n$ with vertices at $o, p_1, \ldots, p_n$ is

$$\frac{1}{n!}|\det(p_{ij})|, \tag{0.13}$$

as in [85, p. 117]. We have the formula

$$\lambda_n(P) = \frac{1}{n} z \lambda_{n-1}(Q) \tag{0.14}$$

for the volume of a pyramid or bounded cone P with base Q and height (the distance from $\operatorname{aff} Q$ to the apex) z. This is easily obtained by integration and induction, as in [52, 9.12.4.4] for the simplex; Dehn's solution of Hilbert's third problem indicates that some form of limit argument is required (see the discussion in [53, 12.2.5.2], for example).

We occasionally need other Borel measures in $\mathbb{E}^n$ or S^{n-1}. A signed Borel measure μ in S^{n-1} is called *even* (or *odd*) if $\mu(-E) = \mu(E)$ (or $\mu(-E) = -\mu(E)$, respectively), for all Borel sets E.

Let μ be a measure in $\mathbb{E}^n$ and E a bounded set in $\mathbb{E}^n$ of finite positive μ-measure. The *centroid* of E with respect to μ is the point

$$c = \frac{1}{\mu(E)} \int_E x \, d\mu(x). \tag{0.15}$$

The centroid of E is contained in $\operatorname{conv} E$; see [83, Section 6, p. 9].

There is another measure that is extremely important in geometry, and it occurs in this fashion. It is sometimes essential to be able to measure the size of a set of lines or planes, or to integrate a function defined on a set of lines or planes. We only need to do this for sets of subspaces, that is, lines and planes containing the origin, or generally for subsets of $\mathcal{G}(n, k)$. Moreover, our measure should be compatible with the appropriate geometric transformations, so that, for example, the measure of a subset E of $\mathcal{G}(n, k)$ should equal the measure of the set obtained by applying the same rotation about the origin to each member of E. For $k = 1$ (or $k = n - 1$), this is easy: Just identify each 1-dimensional subspace (or $(n-1)$-dimensional subspace) S with the corresponding antipodal pair of points $\pm u$ in S^{n-1} such that the vector u is parallel to S (or orthogonal to S, respectively), and then use the measure λ_{n-1} in S^{n-1}. For $1 < k < n - 1$, however, one needs a new measure, which can be defined by the following general process.

Let X be a locally compact topological group. Then there is a nonzero Borel-regular measure μ in X that is also invariant under left translations by elements of X. This measure μ is called the *Haar measure* in X; it is unique up to multiplication by a constant, and is finite if X is compact. A detailed proof of its existence and uniqueness is given in the texts of Cohn [168, Chapter 9] and Munroe [639, Section 17], for example. However, for the special case of most interest here, this can be avoided. A clever direct construction due to Schneider and Weil [746, Satz 1.2.4, p. 21] shows that there is a Haar measure ν_n in the compact group SO_n, normalized so that $\nu_n(SO_n) = 1$. Let $S \in \mathcal{G}(n,k)$, and let $f_k: SO_n \to \mathcal{G}(n,k)$ be defined by $f_k(\phi) = \phi S$ for each $\phi \in SO_n$. This map is surjective; the usual topology for $\mathcal{G}(n,k)$ is the finest topology for which f_k is continuous, and with this topology $\mathcal{G}(n,k)$ is a compact space. If $E \subset \mathcal{G}(n,k)$ is a Borel set, define

$$\mu_{n,k}(E) = \nu_n\big(f_k^{-1}(E)\big).$$

Then $\mu_{n,k}$, also referred to as the Haar measure in $\mathcal{G}(n,k)$, is the measure we need; note that it is normalized so that

$$\mu_{n,k}\big(\mathcal{G}(n,k)\big) = 1,$$

a fact we shall use several times without special comment. When integrating over $\mathcal{G}(n,k)$, we shall abbreviate $d\mu_{n,k}(S)$ by dS, and with this will have no further use for the symbol $\mu_{n,k}$.

Finally, we need a few more definitions. Let X be any set, and μ a measure in X. If $p \geq 1$ is a real number, $L^p(X)$ denotes the set of μ-measurable extended real-valued functions on X such that $\int_X |f(x)|^p \, d\mu(x) < \infty$, and $\|f\|_p = \big(\int_X |f(x)|^p \, d\mu(x)\big)^{1/p}$ is the L^p *norm*. Also, $L^\infty(X)$ is the space of essentially bounded μ-measurable functions on X, with the L^∞ *norm* given by $\|f\|_\infty = \operatorname{ess\,sup}|f|$. (The function f is essentially bounded if it is equal μ-almost everywhere to a bounded function g. The *essential supremum* ess sup f of f is the infimum of the suprema of such g, and the *essential infimum* ess inf f of f is the supremum of the infima of such g. For continuous functions, the essential supremum and infimum reduce to the ordinary supremum and infimum, respectively.)

A sequence $\{\mu_n\}$ of finite Borel measures in a metric space X is said to *converge weakly* to the finite Borel measure μ in X if

$$\int_X f(x)\, d\mu_n(x) \to \int_X f(x)\, d\mu(x), \tag{0.16}$$

for each bounded $f \in C(X)$.

0.6. The support function

Perhaps the most widely applicable function connected with the study of convex sets is the support function, and the purpose of this section is to gather together some of its properties.

If K is a nonempty compact convex set in $\mathbb{E}^n$, the *support function* h_K of K is defined by

$$h_K(x) = \max\{x \cdot y : y \in K\}, \tag{0.17}$$

for $x \in \mathbb{E}^n$. From this definition it follows that if K_1 and K_2 are compact convex sets, then $K_1 \subset K_2$ if and only if $h_{K_1} \leq h_{K_2}$, and this implies that a compact convex set is determined by its support function.

If $u \in S^{n-1}$, then

$$H_u = \{x : x \cdot u = h_K(u)\} \tag{0.18}$$

is the supporting hyperplane to K with outer normal vector u. The support function $h_K(u)$ at a *unit vector* u gives the *signed* distance from o to H_u; see Figure 0.2.

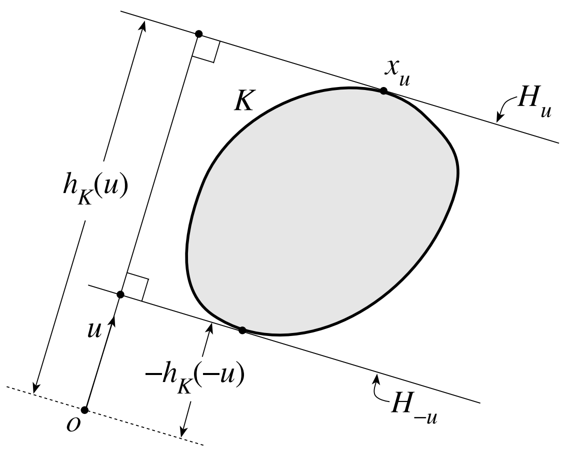


Figure 0.2. The support function.

As a function on $\mathbb{E}^n$, the support function is positively homogeneous, that is,

$$h_K(cx) = ch_K(x) \text{ for } c \geq 0, \tag{0.19}$$

and subadditive, that is,

$$h_K(x + y) \leq h_K(x) + h_K(y). \tag{0.20}$$

See [83, Section 15, p. 26]; a function having both these properties is called *sublinear*. An important result is that the converse is true: Every sublinear function on $\mathbb{E}^n$ is the support function of a unique compact convex set. See [737, Theorem 1.7.1] for three proofs. Every sublinear function is convex, and every convex function is continuous on the interior of its domain (cf. [499, Theorem 30.2] or [737, Theorem 1.5.1]); therefore the support function is convex and continuous on $\mathbb{E}^n$.

Equation (0.19) usually permits us to work with the restriction of h_K to the unit sphere, as we almost always do in this book, without special comment.

The support function has some other properties making it of fundamental importance in convexity. It is immediate from the definition that a compact convex set K is centered if and only if $h_K(u) = h_K(-u)$ for all $u \in S^{n-1}$. It is an easy exercise to show that if S is a subspace, then

$$h_{K|S}(u) = h_K(u), \tag{0.21}$$

for all $u \in S^{n-1} \cap S$. The support function also gives a convenient expression for the Hausdorff distance between two compact convex sets K_1 and K_2, namely, that

$$\delta(K_1, K_2) = \sup_{u \in S^{n-1}} |h_{K_1}(u) - h_{K_2}(u)| = \|h_{K_1} - h_{K_2}\|_\infty. \tag{0.22}$$

The simple proof is given in [737, Theorem 1.8.11].

If K_i is a compact convex set in $\mathbb{E}^n$, and $t_i \geq 0$, $1 \leq i \leq m$, then the vector sum

$$t_1K_1 + \cdots + t_mK_m = \{t_1x_1 + \cdots + t_mx_m : x_i \in K_i\}$$

is also called a *Minkowski linear combination*. The addition and scalar multiplication are also called *Minkowski addition* and *Minkowski scalar multiplication*. It is easily shown that this Minkowski linear combination is itself a compact convex set.

If $t_1, t_2 \geq 0$, then

$$h_{t_1K_1+t_2K_2}(x) = t_1h_{K_1}(x) + t_2h_{K_2}(x), \tag{0.23}$$

for all $x \in \mathbb{E}^n$; for the simple proof, see [737, Theorem 1.7.5].

If K is the singleton set $\{x\}$, then $h_K(u) = u \cdot x$, for all $u \in S^{n-1}$. From this and (0.23) we see that if K is a compact convex set, then the support function of a translate of K is given by

$$h_{K+x}(u) = h_K(u) + x \cdot u, \tag{0.24}$$

for all $u \in S^{n-1}$. It also follows that if $K = [x, y]$ is a line segment, then $h_K(u) = \max\{u \cdot x, u \cdot y\}$; in particular, if $v \in S^{n-1}$, then

$$h_{[-v,v]}(u) = |u \cdot v|, \tag{0.25}$$

for all $u \in S^{n-1}$. Now (0.23) implies that when K is the centered cube in $\mathbb{E}^n$, with sides of length 2 and parallel to the coordinate hyperplanes, we have $h_K(u) = \sum_{i=1}^n |u_i|$, for $u = (u_1, \ldots, u_n) \in S^{n-1}$. Generally, one can show that K is a polytope if and only if h_K is piecewise linear.

Let $\phi \in GL_n$. Then, with (0.1),

$$\begin{aligned} h_{\phi K}(x) &= \max\{x \cdot y : y \in \phi K\} \\ &= \max\{x \cdot \phi z : z \in K\} \\ &= \max\{\phi^t x \cdot z : z \in K\}, \end{aligned}$$

so

$$h_{\phi K}(x) = h_K(\phi^t x), \tag{0.26}$$

for all $x \in \mathbb{E}^n$, and

$$h_{\phi K}(u) = h_K(\phi^t u) = \|\phi^t u\| h_K\left(\frac{\phi^t u}{\|\phi^t u\|}\right), \tag{0.27}$$

for all $u \in S^{n-1}$.

Of course, $h_B(u) = 1$, for all $u \in S^{n-1}$; the support function of an ellipsoid can be obtained from this and (0.27).

0.7. Star sets and the radial function

The radial function is dual to the support function introduced in the previous section, but it appears much less frequently in the literature. Whereas it is natural to define the support function for convex sets, the radial function can be defined for the more general star sets. The purpose of this section is to explain the meaning of these terms.

A set L is *star-shaped* at o if every line through o that meets L does so in a (possibly degenerate) line segment. If L is nonempty, compact, and star-shaped at o, its *radial function* ρ_L is defined by

$$\rho_L(x) = \max\{c : cx \in L\}, \tag{0.28}$$

for $x \in \mathbb{E}^n \setminus \{o\}$ such that the line through x and o meets L. It is positively homogeneous of degree -1 that is,

$$\rho_L(cx) = c^{-1}\rho_L(x) \text{ for } c > 0. \tag{0.29}$$

As with the support function, this usually permits us to work with the restriction of ρ_L to the unit sphere, and we shall do this without further comment. We denote the domain of this restriction by D_L and its support by S_L.

Note that our definition of radial function differs in an important way from the usual one, in which the maximum is taken only over nonnegative c. One advantage of the new definition, introduced by Gardner and Volčič [283], is that it mirrors the definition (0.17) of the support function. The radial function $\rho_L(u)$ at a *unit* *vector* u gives the *signed* distance from o to the boundary of L along the line l_u through o parallel to u. See Figure 0.3.

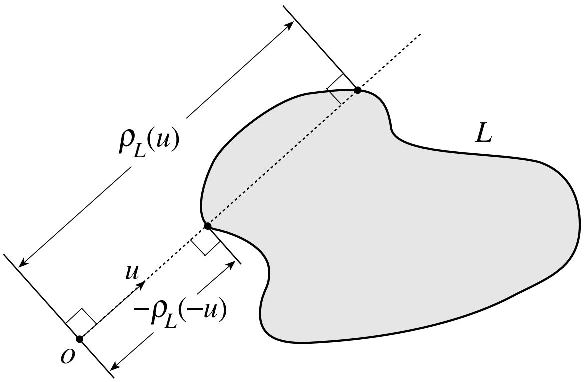


Figure 0.3. The radial function.

By a *star body* we mean a body such that ρ_L, restricted to S_L, is continuous. A *star set* is a set that is a star body in its linear hull.

We warn the reader that there are several definitions of the term "star body" currently in use. One insists that $o \in \operatorname{int} L$, which is clearly more restrictive than the definition here. Another, specifying that ρ_L is continuous, is a viable alternative for bodies containing the origin (see Chapter 8, especially the discussion at the beginning of Section 8.1). Our definition has the advantage that a convex body is always a star body. Sometimes, however, the extra assumption that S_L is centered is required, as in Theorem 7.2.3, for example.

Let L be a star set. If $o \in L$, then D_L coincides with S^{n-1}; otherwise, D_L is smaller. Let L be a star body. It is not difficult to see that since L is compact and regular, both D_L and S_L are also compact and regular in S^{n-1}. Since ρ_L is continuous on S_L, it is a bounded Borel function on D_L.

If L is a star set in $\mathbb{E}^n$, and $S \in \mathcal{G}(n,k)$, then $L \cap S$ need not be a star set, since it may not be regular. However, if $\dim(L \cap S) = k$, then $L \cap S$ is a star set.

If $x_i \in \mathbb{E}^n$, $1 \le i \le m$, then $x_1 \tilde{+} \cdots \tilde{+} x_m$ is defined to be the usual vector sum of the points x_i, if all of them are contained in a line through o, and o otherwise. Let L_i be a star body in $\mathbb{E}^n$ with $o \in L_i$, and $t_i \ge 0$, $1 \le i \le m$; then

$$t_1 L_1 \tilde{+} \cdots \tilde{+} t_m L_m = \{t_1 x_1 \tilde{+} \cdots \tilde{+} t_m x_m : x_i \in L_i\} \tag{0.30}$$

is called a *radial linear combination*. The addition and scalar multiplication are called *radial addition* and *radial scalar multiplication*. (Lutwak [537] adds Minkowski's name to these terms.) Moreover,

$$\rho_{t_1 L_1 \tilde{+} t_2 L_2}(x) = t_1 \rho_{L_1}(x) + t_2 \rho_{L_2}(x), \tag{0.31}$$

for all x.

One can measure distance between star bodies by means of the Hausdorff metric. However, in many respects the *radial metric* $\tilde{\delta}$ is more natural. This is defined by setting

$$\tilde{\delta}(L_1, L_2) = \sup_{u \in S^{n-1}} |\rho_{L_1}(u) - \rho_{L_2}(u)| = \|\rho_{L_1} - \rho_{L_2}\|_\infty, \tag{0.32}$$

for star bodies L_1, L_2 in $\mathbb{E}^n$.

Let $\phi \in GL_n$. Then it follows from the definition of ρ_L that

$$\rho_{\phi L}(x) = \rho_L(\phi^{-1}x), \tag{0.33}$$

for $x \in \mathbb{E}^n \setminus \{o\}$, so

$$\rho_{\phi L}(u) = \rho_L(\phi^{-1}u) = \frac{1}{\|\phi^{-1}u\|}\rho_L\left(\frac{\phi^{-1}u}{\|\phi^{-1}u\|}\right), \tag{0.34}$$

for all $u \in S^{n-1}$.

Many examples of radial functions can be obtained from those of support functions via the important polar relation (0.36).

0.8. Polar duality

Polar duality is an important tool in geometry, and it will be used several times in this book. Though much is known about polar duality, it can be frustrating to search the literature for even the most basic facts, so these are collected together in this section.

If E is an arbitrary nonempty subset of $\mathbb{E}^n$, then the set

$$E^* = \{x : x \cdot y \le 1 \text{ for all } y \in E\} \tag{0.35}$$

is called the *polar set* of E. The polar set is always closed and convex and contains the origin; see [499, p. 142]. Moreover, if K is a convex body and $o \in \operatorname{int} K$, then the same is true of K^*, which we then call the *polar body* of K; see Figure 0.4. In this case $K^{**} = K$ (see [499, p. 142], [595, Section 2.2], or [737, Theorem 1.6.1]).

If K is a convex body in $\mathbb{E}^n$ such that $o \in \operatorname{int} K$, the boundary of K^* can be calculated by the following important relation (see [499, Theorem 29.8] or [737, Remark 1.7.7]):

$$\rho_{K^*}(u) = 1/h_K(u), \tag{0.36}$$

for all $u \in S^{n-1}$.

Suppose that K is a centered convex body in $\mathbb{E}^n$. Then (0.36) and the fact that $K^{**} = K$ show that the reciprocal of ρ_K (sometimes called the *gauge function* of K) is h_{K^*} and is therefore sublinear. In view of the fact that K is centered, the reciprocal of ρ_K actually defines a norm $\|\cdot\|_K$ on $\mathbb{E}^n$, for which K is the unit ball.

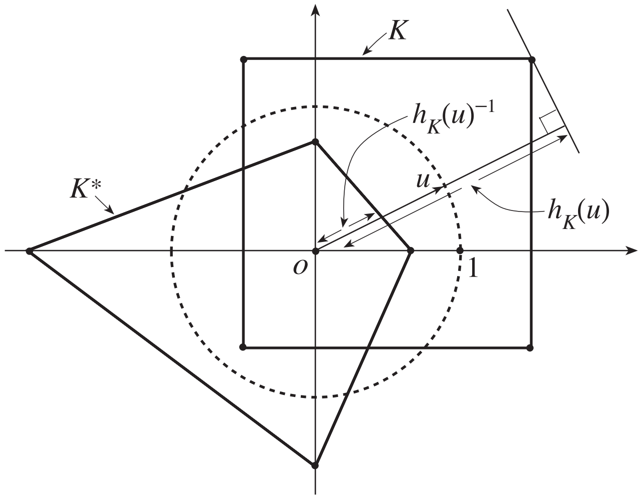


Figure 0.4. The polar body.

The dual of this Banach space is the one whose norm is given by h_K, which has the polar body K^* as its unit ball; see [737, Remark 1.7.8]. This is the source of an intimate and important connection between centered convex bodies and Banach spaces.

The polar set of the single point $\{x\}$, $x \neq o$, is the half-space $\{y : x \cdot y \leq 1\}$, and the polar body of the ball rB is the ball $r^{-1}B$. The following examples of polar bodies of convex polytopes are noted by Grünbaum [367, Section 3.4]. If P is an n-dimensional simplex containing the origin in its interior, then P^* is also. The polar body of the centered unit cube in $\mathbb{E}^n$ is a centered cross-polytope. A centered regular dodecahedron has a centered icosahedron for its polar body. Generally, the polar body of a convex polytope P is also a convex polytope, and polarity provides an inclusion-reversing bijection between the faces of P and the faces of P^* (see [595, Lemma 8, p. 65]).

Suppose that $\phi \in GL_n$, and that K is a convex body in $\mathbb{E}^n$ with $o \in \operatorname{int} K$. Then for each $u \in S^{n-1}$, by (0.27) and (0.34),

$$h_{(\phi K)^*}(u) = \frac{1}{\rho_{\phi K}(u)} = \frac{1}{\rho_K(\phi^{-1}u)} = h_{K^*}(\phi^{-1}u) = h_{\phi^{-t}K^*}(u).$$

Therefore

$$(\phi K)^* = \phi^{-t}K^*. \tag{0.37}$$

It follows immediately from (0.37) that centered convex bodies are similar, or linearly equivalent, if and only if their polar bodies are, respectively.

If E is an n-dimensional ellipsoid in $\mathbb{E}^n$, containing the origin in its interior, then E^* is also. To see this, note first that if $a = (a_1, \dots, a_n) \in \operatorname{int} B$, then by (0.36) and (0.24),

$$\rho_{(B+a)^*}(u) = \frac{1}{h_{(B+a)}(u)} = \frac{1}{1 + a \cdot u},$$

for all $u \in S^{n-1}$. This gives

$$\rho_{(B+a)^*}(u) = 1 - \rho_{(B+a)^*}(u) u \cdot a.$$

If we let $x = (x_1, \dots, x_n) = \rho_{(B+a)^*}(u)u$, the previous equation becomes

$$\sum_{i=1}^{n} x_i^2 = \left(1 - \sum_{i=1}^{n} a_i x_i\right)^2.$$

Since this equation is a quadratic, and since we know that $(B + a)^*$ is convex, it follows that $(B + a)^*$ is an ellipsoid (see [53, Proposition 15.4.7]). Now if $E = \phi(B + a)$, where $a \in \operatorname{int} B$ and $\phi \in GL_n$, then (0.37) implies that $E^* = \phi^{-t}(B + a)^*$ is an ellipsoid.

Polar duality provides a link between sections and projections. Indeed, suppose that K is a convex body in $\mathbb{E}^n$ with $o \in \operatorname{int} K$, and that S is a subspace. Then

$$K^* \cap S = (K|S)^*, \tag{0.38}$$

where the polar operation on the right is taken in S. One can see this by using (0.36) and (0.21) to conclude that, for any $u \in S \cap S^{n-1}$,

$$\rho_{K^* \cap S}(u) = \rho_{K^*}(u) = \frac{1}{h_K(u)} = \frac{1}{h_{K|S}(u)} = \rho_{(K|S)^*}(u);$$

or see [595, Theorem 15, p. 70] for a simple proof using only the definition of a polar body.

Despite (0.38), the use of polar duality in geometric tomography is severely limited by the fact that (as Figure 0.4 suggests) it is not an affine notion, but rather a projective one. To be more specific, consider the following result, proved in [595, Theorem 14, p. 67]. Let K be a convex body in $\mathbb{E}^n$ with $o \in \operatorname{int} K$, and let ϕ be a nonsingular projective transformation of $\mathbb{E}^n$, permissible for K, such that $o \in \operatorname{int} \phi K$. Then there is a nonsingular projective transformation ψ, permissible for K^*, such that $(\phi K)^* = \psi K^*$.

0.9. Differentiability properties

A real-valued function on an open subset U of $\mathbb{E}^n$ is said to be *of class C^k* if it is k-times continuously differentiable, that is, all partial derivatives of order k exist and are continuous. The class of such functions is signified by $C^k(U)$. The class $C^\infty(U)$ consists of those real-valued functions belonging to $C^k(U)$ for all $k \in \mathbb{N}$. A real-valued function f on U is *real analytic* if its Taylor series exists and converges to $f(x)$ at each $x \in U$. (See, for example, [570, Section 6.8].)

If $n \in \mathbb{N}$ and $1 \leq i \leq n$, let π_i be the real-valued function on $\mathbb{E}^n$ defined by $\pi_i(x) = x_i$, where $x = (x_1, \ldots, x_n)$. Suppose that f is a function from an open subset U of $\mathbb{E}^n$ into $\mathbb{E}^m$. Then we say f is of class C^k if each map $\pi_i \circ f$, $1 \leq i \leq m$, is in $C^k(U)$. Functions from an open subset of $\mathbb{E}^n$ into $\mathbb{E}^m$ that are of class C^∞ or real analytic are defined analogously.

Sometimes we want to speak, for example, about a function belonging to $C^\infty(S^{n-1})$, or the boundary of a convex body being of class C^2. Basically, the meaning of such terms is inherited from those defined in the previous paragraph, but precise definitions take a little work. These can be found in several books on differential geometry, but for the convenience of the reader we present them here.

A subset M of $\mathbb{E}^n$ is called an *m-dimensional submanifold of* $\mathbb{E}^n$ *of class* C^k if there is an *atlas* for M of class C^k. An atlas for M of class C^k is a family of pairs (U_r, f_r), called *charts*, such that

(i) each U_r is an open subset of M, and $\cup_r U_r = M$;

(ii) each f_r is a homeomorphism of U_r onto an open subset of $\mathbb{E}^m$;

(iii) if $U_r \cap U_s \neq \emptyset$, the map $f_s \circ f_r^{-1}$, from the open subset $f_r(U_r \cap U_s)$ of $\mathbb{E}^m$ into $\mathbb{E}^m$, is of class C^k.

Again, an *m-dimensional submanifold of* $\mathbb{E}^n$ *of class* C^∞ and a *real-analytic m-dimensional submanifold of* $\mathbb{E}^n$ are defined analogously.

The unit sphere S^2 is an example of a real-analytic 2-dimensional submanifold of $\mathbb{E}^3$. The reason for the somewhat technical definition given is that one cannot map the whole of S^2 onto an open subset of the plane in the appropriate way, so one has to use several charts; the sets U_r are patches on S^2 which cover it and which can be mapped onto open subsets of the plane. (In fact, the term "patch" is often used instead of "chart.") A picture of such patches covering a surface is one of many figures generated with Mathematica by Gray [330, p. 219].

We say that a convex body K is *of class* C^k or *of class* C^∞ if bd K is of class C^k (or C^∞, respectively) as a submanifold of $\mathbb{E}^n$.

Now let M be an m-dimensional submanifold of $\mathbb{E}^n$ of class C^l, and suppose that $k \leq l$. The real-valued function f on M is of class C^k if, for every chart (U_r, f_r) in an atlas for M, the real-valued function $f \circ f_r^{-1}$ on the open subset $f_r(U_r)$ of $\mathbb{E}^m$ is of class C^k. The class of such functions is denoted by $C^k(M)$. The real-valued functions on M of class C^∞, or real-analytic ones, are defined analogously, and the former class is signified by $C^\infty(M)$.

When E is an appropriate subset of $\mathbb{E}^n$, $C_e^k(E)$ and $C_e^\infty(E)$ denote the even functions in the classes defined earlier.

Occasionally we adopt the common practice of calling a function or body "sufficiently smooth." This simply means that it belongs to C^k for a sufficiently large k.

If U is an open subset of the reals, the notation $f \in C^{k+\varepsilon}(U)$, $0 < \varepsilon \leq 1$, means that $f \in C^k(U)$ and the kth derivative $f^{(k)}$ of f satisfies the following *Hölder condition*:

$$|f^{(k)}(x) - f^{(k)}(y)| \leq c\,|x-y|^{\varepsilon}, \tag{0.39}$$

for some $c \geq 0$ and all $x, y \in U$. This allows some of the previous definitions to have meaning even when k is an arbitrary nonnegative real number.

In the early literature on convexity, it was standard procedure for authors to assume any convenient order of smoothness of the boundary of a convex body, often without explicit comment. As the years went by, it happened again and again that examples were discovered of convex bodies with surprisingly complicated boundary structure. Luckily, one can learn basic convexity without spending too much time on this topic, and in the advanced theory of convexity, Aleksandrov's magnificent theory of area measures (see Section A.2) often allows one to proceed without any special boundary assumptions.

Sooner or later, however, one has to make such assumptions. Unfortunately, there are many pitfalls and highly nonintuitive phenomena. For example, in [436], Kiselman has shown the existence of a convex body in $\mathbb{E}^3$ having a real-analytic boundary surface, though the boundary of its projection on some plane is only $C^{2+\varepsilon}$ for some $\varepsilon > 0$; and in [437], he proves that the boundary of a Minkowski sum of two planar convex bodies with real-analytic boundaries is $C^{20/3}$, and that this result is the best possible! Moreover, basic results and even definitions tend to be involved and require some knowledge of differential geometry. Until recently, there was no adequate treatment available, but Chapter 2 of Schneider's book [737] now provides a lucid and extremely valuable guide, plundered for the following summary.

Let K be a convex body in $\mathbb{E}^n$. The support function h_K is differentiable on $\mathbb{E}^n \setminus \{o\}$ if and only if h_K is C^1, and also if and only if K is strictly convex; see [737, p. 107]. In this case $H_u \cap K$ (where H_u is defined by (0.18)) is a single point x_u for each $u \in S^{n-1}$, and

$$x_u = \operatorname{grad} h_K(u),$$

where grad denotes gradient (see [737, Corollary 1.7.3]). From this, the boundary of K can be computed. For $n = 2$, we can do this directly, as follows. If h_K is differentiable at $u = (\cos\theta, \sin\theta) \in S^1$, then $h_K(u) = x_u \cdot u$. Regarding h_K as a function of θ, we get

$$h'_K(\theta) = x'_u \cdot u + u' \cdot x_u,$$

where the primes denote differentiation with respect to θ. Now x'_u is parallel to the tangent to K at x_u, so $x'_u \cdot u = 0$. Also, u' is orthogonal to u. It follows that $|h'_K(\theta)|$ is the distance from x_u to the foot of the perpendicular from o to H_u, and that

$$x_u = \big(h_K(\theta)\cos\theta - h_K'(\theta)\sin\theta, h_K(\theta)\sin\theta + h_K'(\theta)\cos\theta\big). \qquad (0.40)$$

The terms “convex body of class C^k (or C^∞)” and “smooth” have already been explained. In [737, p. 104], it is noted that K is smooth if and only if it is of class C^1. The proof given there requires several basic differentiability properties of convex functions.

Let K be smooth and $x \in \operatorname{bd} K$. Suppose that u is the outer unit normal vector to K at x. The *Gauss map* g from $\operatorname{bd} K$ to S^{n-1} is defined by $g(x) = u$; it is continuous, and a homeomorphism if K is smooth and strictly convex (see [737, p. 78]).

The *tangent space* of K at x is the translate $H_u - x = u^\perp$ of the supporting hyperplane to K with outer normal vector u. Suppose now that K is of class C^2. Then g is C^1. The differential $W_x = dg_x$ of the Gauss map is a linear map from this tangent space to itself, called the *Weingarten map*. The eigenvalues of W_x are called the *principal curvatures* of K at x. (In $\mathbb{E}^3$, the principal curvatures at a point in $\operatorname{bd} K$ give the maximal and minimal bending of $\operatorname{bd} K$ at the point.) The principal curvatures are nonnegative (see [737, pp. 104–6]). Their product is the *Gauss curvature* (or Gauss–Kronecker curvature) of K at x.

The Gauss curvature at a point on an $(n-1)$-dimensional submanifold of $\mathbb{E}^n$ or *hypersurface* of class C^2 can be defined similarly, as in [452, Chapter 7, Section 5]. Then a compact hypersurface forms the boundary of a strictly convex body if and only if the Gauss curvature at each of its points is positive; see [452, Theorem 5.6].

If $r = r(\theta)$ is a planar C^2 curve, we have the well-known formula

$$\frac{2(r')^2 - rr'' + r^2}{\big((r')^2 + r^2\big)^{3/2}} \qquad (0.41)$$

for its curvature in polar coordinates.

If $k \geq 2$, we say that K is *of class* C^k_+ (or C^∞_+) if K is of class C^k (or C^∞, respectively) and the Gauss curvature of K at each x is positive.

Suppose that $h_K \in C^2$. Since this implies that K is strictly convex, the *reverse spherical image map* from S^{n-1} to $\operatorname{bd} K$, taking u to x_u, is defined. Furthermore, its differential $\overline{W}_u = dx_u$ is also defined; this linear map from the tangent space $u^\perp$ of S^{n-1} at u to itself is called the *reverse Weingarten map*. The eigenvalues of $\overline{W}_u$ are called the *principal radii of curvature* of K at $u \in S^{n-1}$. We denote them by $R_i(u)$, $1 \leq i \leq n-1$, where the labeling ranks them by magnitude. (See [737, pp. 107–8], where the notation $r_i(u)$ is used instead. The corresponding eigenvectors are called *principal curvature directions*.) They are also the nonzero eigenvalues of the second differential of h_K at u, by [737, Corollary 2.5.2]. If K is of class C^2_+, they are also the eigenvalues of the inverse Weingarten map

$W_{x_u}^{-1} = dg_{x_u}^{-1}$, and coincide with the reciprocals of the principal curvatures of K at $g^{-1}(u)$.

A *principal curve* in bd K is a curve whose tangent vectors point in a principal curvature direction. Figure 3.9 shows some principal curves in the boundary of an ellipsoid.

Notice that K is of class C_+^2 if and only if K is of class C^2 and all the principal curvatures of K are everywhere positive, and also (since K must be smooth, by the preceding remarks) if and only if all the principal radii of curvature are everywhere finite and positive.

It is proved in [737, pp. 106–11] that K is of class C_+^2 if and only if $h_K \in C^2$ and K has positive finite principal radii of curvature, or equivalently, if and only if $h_K \in C^2$ and the Gauss curvature of K exists and is positive everywhere. The existence of the Gauss curvature is necessary, since it is possible that $h_K \in C^2$ and the Gauss curvature of K is positive everywhere it exists, yet K is not even smooth. (In $\mathbb{E}^2$, for example, K may have a corner point x so $h_K(u) = h_{\{x\}}(u) = x \cdot u$ is linear for $u \in S^{n-1}$ in a neighborhood of $x/\|x\|$.) The positivity of the Gauss curvature is also essential; Hartman and Wintner [383, p. 480] have shown that h_K is not necessarily C^2 even if the boundary of K is real analytic.

Let $F_K^{(i)}$ be defined by

$$\binom{n-1}{i} F_K^{(i)} = \sum_{1 \le j_1 < \cdots < j_i \le n-1} R_{j_1} \cdots R_{j_i}, \tag{0.42}$$

where the right-hand side is the ith elementary symmetric function of the principal radii of curvature of K. (In [737] the notation s_i is used instead of $F_K^{(i)}$.) Then (see [737, Corollary 2.5.3] or [83, Section 38]) $F_K^{(i)}(u)$ is the sum of the principal minors of order i of the Hessian matrix of h_K at u.

Using a result of Aleksandrov, it is possible to define, for almost all $u \in S^{n-1}$, the principal radii of curvature $R_i(u)$ of a general convex body K in $\mathbb{E}^n$ as the eigenvalues, corresponding to eigenvectors orthogonal to u, of the second differential of h_K at $u \in S^{n-1}$. Thus each $F_K^{(i)}$ can similarly be defined for almost all $u \in S^{n-1}$. Each R_i, and hence each $F_K^{(i)}$, is then integrable on S^{n-1} with respect to spherical Lebesgue measure; see [737, p. 118].

The *Laplacian* $\triangle f$ of a function f on an open subset of $\mathbb{E}^n$ is defined, as usual, by

$$\triangle f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}.$$

The *Laplace–Beltrami operator* $\triangle_S$ on S^{n-1} can be defined as follows. If f is a function on S^{n-1}, it can be extended to a function g on $\mathbb{E}^n \setminus \{o\}$ by letting $g(x) = f(x/\|x\|)$ for each $x \neq o$. Then $\triangle_S f$ is defined to be the restriction of $\triangle g$ to S^{n-1}.

If we set $i = 1$, we get (cf. [737, p. 110])

$$(n-1)F_K^{(1)} = R_1 + \cdots + R_{n-1} = (n-1)h_K + \triangle_S h_K.$$

In particular, if $n = 2$ and $u = (\cos\theta, \sin\theta)$, we obtain, as in [737, p. 110, eq. (2.5.22)],

$$R_1(u) = h_K(\theta) + h_K''(\theta). \tag{0.43}$$

In Section 0.4 it was noted that a convex body can be approximated arbitrarily closely in the Hausdorff metric by smooth convex bodies. Sometimes it is useful to have even stronger approximation theorems. For example, it follows from [737, pp. 158–60] that the class of C_+^∞ convex bodies is dense in $\mathcal{K}^n$, and further that each centrally symmetric member of $\mathcal{K}^n$ can be approximated by centrally symmetric C_+^∞ convex bodies.

1

Parallel X-rays of planar convex bodies

In this chapter our goal is to investigate the tomography of convex bodies in the plane. The requisite concepts of an X-ray and Steiner symmetral of a planar convex body are introduced in such a way that no knowledge of measure theory or Lebesgue integration is necessary. Furthermore, the reader can absorb the new ideas while avoiding the technicalities of higher-dimensional spaces. More general definitions are postponed until Chapter 2. (Occasional reference is made to these, but this is merely for cross-reference.) Granted some (but by no means all) of the background material in the first four sections of Chapter 0, and apart from references to a couple of auxiliary facts, the chapter is self-contained.

An X-ray of a convex body gives the lengths of all the chords of the body parallel to the direction of the X-ray. Corollary 1.2.12 states that there are four directions such that every convex body is determined, among all convex bodies, by its X-rays in these directions. Given a convex body, Theorem 1.2.21 says that there are three directions allowing the body to be distinguished from all others – "verified" – by the corresponding X-rays. A practical method by which every convex polygon can be "successively determined" by three X-rays, the direction of each depending only on the previous X-rays, is provided by Theorem 1.2.23.

1.1. What is an X-ray?

We all know that dense material such as bone or teeth will show as light areas on a medical X-ray in a doctor's viewing box, while darker regions correspond to other less dense tissue. Each beam of the X-ray travels along a straight line, and its intensity after traversing the body depends on how much material it has passed through; high intensities result in a darker point on the X-ray picture, and low intensities show as a lighter point. If the beams in the X-ray are all parallel, then the X-ray picture contains information about the amount of material in the body lying on each straight line parallel to the direction of the X-ray.