PURE & APPLIED MATHEMATICS
A WILEY-INTERSCIENCE SERIES OF
TEXTS, MONOGRAPHS & TRACTS

STEVEN R. LAY

CONVEX SETS AND THEIR APPLICATIONS

CONVEX SETS AND THEIR APPLICATIONS

STEVEN R. LAY
Department of Mathematics
Aurora College

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS

New York · Chichester · Brisbane · Toronto · Singapore

Library of Congress Cataloging in Publication Data:

Lay, Steven R., 1944–
Convex sets and their applications.

(Pure and applied mathematics, ISSN 0079-8185)
"A Wiley-Interscience publication."
Bibliography: p.
Includes index.
1. Convex sets. I. Title. II. Series: Pure and applied mathematics (John Wiley & Sons)
QA640.L38 511.3'2 81-19738
ISBN 0-471-09584-2 AACR2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To

Frederick A. Valentine,

who first introduced me to the beauty of a convex set

PREFACE

The study of convex sets is a fairly recent development in the history of mathematics. While a few important results date from the late 1800s, the first systematic study was the German book *Theorie der konvexen Körper* by Bonnesen and Fenchel in 1934. In the 1940s and 1950s many useful applications of convex sets were discovered, particularly in the field of optimization. The importance of these applications has in turn sparked a renewed interest in the theory of convex sets.

The relative youth of convexity has both good and bad implications for the interested student. Unlike most older branches of mathematics, there is not a vast body of background material that must be mastered before the student can reach significant unsolved problems. Indeed, even high school students can understand some of the basic results. Certainly upper level undergraduates have all the tools necessary to explore the properties of convex sets. This is good! Unfortunately, however, since the subject is so new, it has not yet "filtered down" beneath the graduate level in any comprehensive way. To be sure, there are several undergraduate books that devote two or three chapters to some aspects of convex sets, but there is no text at this level which has convex sets and their applications as its unifying theme. It is this void that the present book seeks to fill.

The mathematical prerequisites for our study are twofold: linear algebra and basic point-set topology. The linear algebra is important because we will be studying convex subsets of n-dimensional Euclidean space. The reader should be familiar with vectors and their inner product. The topological concepts we will encounter most frequently are open sets, closed sets, and compactness. Occasionally we will need to refer to continuity, convergence of sequences, or connectedness. These topics are typically covered in courses such as Advanced Calculus or Introductory Topology.

While the intent of this book is to introduce the reader to the broad scope of convexity, it certainly cannot hope to be exhaustive. The selection of topics has been influenced by the following goals: The material should be accessible to students with the aforementioned background; the material should lead the student to open questions and unsolved problems; and the material should highlight diverse applications. The degree to which the latter two goals are attained varies from chapter to chapter. For example, Chapter 4 on Kirchberger-type theorems includes few applications, but it does advance to the very edge of current mathematical research. On the other hand, the investigation of polytopes has progressed quite rapidly and so Chapter 8 can only be considered a brief introduction to the subject. Chapter 10 on optimization includes no unsolved problems, but has a major emphasis on applications.

For the past six years the material in this book has been used as a text in an upper-level undergraduate course in convexity. Since there is more material than can be covered in a three-unit semester course, the emphasis has varied from year to year depending on which topics were included. As a result of this flexibility, the material can be used as a geometry text for secondary teachers, as a resource for selected topics in applications, and as a bridge to higher mathematics for those continuing on to graduate school.

Throughout the book, the material is presented with the undergraduate student in mind. This is not intended to imply a lack of rigor, but rather to explain the inclusion of more examples and elementary exercises than would be typical in a graduate text. It is recognized that a course in convexity is more frequently found at the graduate level, but our attempt to make the material as clear and accessible as possible should not detract from its usefulness at this higher level as well. The exercises include not only computational problems, but also proofs requiring various levels of sophistication. The answers to many of the exercises are found at the back of the book, as are hints and references to the literature.

The first three chapters of the book form the foundation for all that follows. We begin in Chapter 1 by reviewing the fundamentals of linear algebra and topology. This enables us to establish our notation and leads naturally into the basic definition and properties of convex sets. Chapter 2 develops the important relationships between hyperplanes and convex sets, and this is applied in the following chapter to the theorems of Helly and Kirchberger.

Beginning with Chapter 4, the chapters become relatively independent of each other. Each chapter develops a particular aspect or application of convex sets. In Chapter 4 we look at several Kirchberger-type theorems in which the separating hyperplanes have been replaced by other geometrical figures. The results presented in this chapter are very recent, having only appeared in the literature in the last couple of years. Virtually any question that is posed beyond the scope of the text opens the door to an unsolved problem. (The answers to these will not be found in the back of the book!)

Chapter 5 investigates a number of topics which have had historical significance in the plane. After looking at sets of constant width and universal covers, we solve the classical isoperimetric problem under the assumption that a set of maximal area with fixed perimeter exists. In Chapter 6 we develop some of the tools necessary to prove the existence of such an extremal set.

There are many different ways to characterize convex sets in terms of local properties, and these are presented in Chapter 7. Chapter 8 introduces us to the basic properties of polytopes and gives special attention to how cubes,

CONTENTS

LIST OF NOTATION

R^n	n-dimensional linear space	1
E^n	n-dimensional Euclidean space	1
θ	the origin	2
$\varnothing$	the empty set	2
iff	if and only if	2
$\equiv$	equality (by definition)	2
R	the set of real numbers	2
$\langle x, y\rangle$	the inner product of x and y	2
$\Vert x\Vert$	the norm of x	2
$d(x, y)$	the distance from x to y	3
$B(x, \delta)$	the open ball of radius δ centered at x	4
$\sim S$	the complement of S	4
int S	the interior of S	5
cl S	the closure of S	5
bd S	the boundary of S	8
$\overline{xy}$	the line segment joining x and y	10
dim S	the dimension of S	12
relint S	the relative interior of S	12
relint $\overline{xy}$	the line segment $\overline{xy}$ without the endpoints	12
conv S	the convex hull of S	16
aff S	the affine hull of S	16
pos S	the positive hull of S	23
$[f: \alpha]$	$\{x \in \mathsf{E}^n: f(x) = \alpha\}$	27
$f(A) \leqslant \alpha$	$f(x) \leqslant \alpha$ for each $x \in A$	33
A_δ	the δ-parallel body to A	94
$D(A, B)$	the distance between A and B	95
$\mathcal{C}$	the set of all nonempty compact convex subsets of E^n	96

k-face	a k-dimensional face	116
$f_k(P)$	the number of k-faces of a polytope P	123
dc M	the dual cone of M	147
P_I and P_{II}	the first and second players in a matrix game	154
$E(x, y)$	the expected payoff for strategies x and y	157
$v(x)$	the value of a strategy x	158
v_I and v_{II}	the value of a game to P_I and P_{II}, respectively	158
epi f	the epigraph of f	199

1

FUNDAMENTALS

A study of convex sets can be undertaken in a variety of settings. The only necessary requirement is the presence of a linear structure, namely, a linear (or vector) space. Indeed, many interesting and useful results can be proved in this general context. At the other extreme are those concepts which seem to find their fulfillment in the Cartesian plane or three-space. We will chart a course somewhere in between these extremes and pursue our study of convex sets in n-dimensional Euclidean space. This setting is broad enough to include many of the important applications of convex sets, yet narrow enough to simplify many of the proofs. Very often the greatest difficulty in extending a result to n-dimensional spaces is encountered in going from two to three dimensions, and here our intuition is more reliable and we are aided by the ability to draw pictures. Whenever a new concept or result is presented, one should immediately construct examples in two and three dimensions to get a better "feeling" for the ideas involved.

SECTION 1. LINEAR ALGEBRA AND TOPOLOGY

The collection of all ordered n-tuples of real numbers (for $n = 1, 2, 3, \dots$) can be made into the real linear space R^n by defining $(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) \equiv (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$ and $\lambda(\alpha_1, \dots, \alpha_n) \equiv (\lambda\alpha_1, \dots, \lambda\alpha_n)$ for any n-tuples $(\alpha_1, \dots, \alpha_n)$ and $(\beta_1, \dots, \beta_n)$ and any real number λ. We define the inner product $\langle x, y \rangle$ of $x \equiv (\alpha_1, \dots, \alpha_n)$ and $y \equiv (\beta_1, \dots, \beta_n)$ to be the real number $\langle x, y \rangle \equiv \Sigma_{i=1}^{n} \alpha_i \beta_i$. The linear space R^n together with the inner product just defined is called n-dimensional Euclidean space, and is denoted by E^n. The n-tuples in E^n are referred to as points or vectors, interchangeably.

Throughout this book we will be dealing with subsets of n-dimensional Euclidean space unless otherwise indicated. If the particular dimension is important, it will be specified. Otherwise, the reader may assume that the context is E^n.

We will usually use lowercase letters such as x, y, z to denote points (or vectors) in E^n. Occasionally, x and y will be used as real variables when giving examples in E^2. For example we may write the linear equation $2x + 3y = 4$. The context should make the usage clear.

Capitals like A, B, C will denote subsets of E^n, and Greek letters such as α, β, δ will denote real scalars. The origin $(0, 0, \ldots, 0)$ will be denoted by θ, and the empty set by $\varnothing$. We will write $A \subset B$ to denote that A is a subset (proper or improper) of B. The shorthand "iff" means "if and only if," and the three-barred equal sign, $\equiv$, is used in defining a new point, set, or function. The set of real numbers will be denoted by R.

Our first theorem is a fundamental result from linear algebra, and its proof is left as an exercise.

1.1. Theorem. The inner product of two vectors in E^n has the following properties for all x, y, z in E^n:

(a) $\langle x, x\rangle \geqslant 0$ and $\langle x, x\rangle = 0$ iff $x = \theta$.
(b) $\langle x, y\rangle = \langle y, x\rangle$.
(c) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.
(d) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for every real α.

1.2. Definition. If $\langle x, y\rangle = 0$, then x and y are said to be **orthogonal** to each other.

By using the inner product in E^n we can talk about the "size" of a vector. Specifically, we define the norm of a vector as follows:

1.3. Definition. The **norm** of a vector x (denoted by $\|x\|$) is given by $\|x\| \equiv \langle x, x\rangle^{1/2}$. If $\|x\| = 1$, then x is called a **unit vector**.

The following properties of the norm are very useful:

1.4. Theorem. For all vectors x and y and real scalar α, the following hold:

(a) $\|x\| > 0$ if $x \neq \theta$, and $\|\theta\| = 0$.
(b) $\|\alpha x\| = |\alpha|\, \|x\|$.
(c) $\|x + y\| \leqslant \|x\| + \|y\|$.
(d) $\langle x, y\rangle = \|x\|\, \|y\| \cos\gamma$, where γ is the angle between the vectors x and y.

PROOF. Parts (a) and (b) follow directly from the definition, and part (c) follows from the Schwarz inequality:

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i\right)^2 \leqslant \left(\sum_{i=1}^{n} \alpha_i^2\right)\left(\sum_{i=1}^{n} \beta_i^2\right).$$

Part (d) follows from the law of cosines. The details are left as an exercise. ∎

By using the preceding norm we can define the distance between two points as follows:

1.5. Definition. If $x, y \in \mathsf{E}^n$, then the **distance from *x* to *y***, denoted by $d(x, y)$, is given by

$$d(x, y) \equiv \|x - y\|.$$

In terms of the inner product we have

$$d(x, y) = \langle x - y, x - y \rangle^{1/2},$$

and in terms of coordinates we have

$$d(x, y) = \left[\sum_{i=1}^{n} (\alpha_i - \beta_i)^2 \right]^{1/2},$$

where $x = (\alpha_1, \ldots, \alpha_n)$ and $y = (\beta_1, \ldots, \beta_n)$.

Examples

1. If $n = 1$, then the distance from x to y is just

$$[(x - y)^2]^{1/2} = |x - y|.$$

2. If $n = 2$, then the distance from x to y is given by the usual formula resulting from the Pythagorean Theorem. (See Figure 1.1.) The length of side a is $|\alpha_1 - \beta_1|$. The length of side b is $|\alpha_2 - \beta_2|$. Thus the length of the hypotenuse c is $d(x, y) \equiv \sqrt{(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2}$.

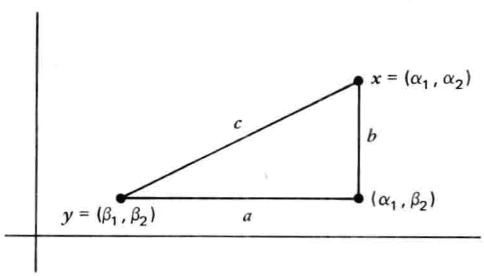

Figure 1.1.

1.6. Theorem. The distance function d has the following properties for all x, y, z in E^n:

(a) $d(x, y) \geqslant 0$ and $d(x, y) = 0$ iff $x = y$.
(b) $d(x, y) = d(y, x)$.
(c) $d(x, y) \leqslant d(x, z) + d(z, y)$.
(d) $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ for every real λ.
(e) $d(x + z, y + z) = d(x, y)$.

PROOF. Parts (a) and (b) follow directly from Theorem 1.1. Part (c) is called the Triangle Inequality and says intuitively that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides. (See Figure 1.2.) The proof of parts (c) and (d) follow from Theorem 1.4. Part (e) follows directly from the definition. ■

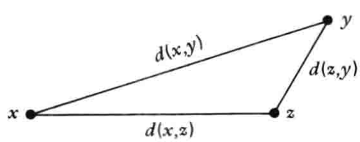

Figure 1.2.

Using the distance function we can define a topology for E^n just as we would for any other metric space.

1.7. Definition. For any $x \in \mathsf{E}^n$ and $\delta > 0$, the **open ball** $B(x, \delta)$ with center x and radius δ is given by

$$B(x, \delta) \equiv \{y \in \mathsf{E}^n : d(x, y) < \delta\}.$$

1.8. Definition. A point x is an **interior point** of the set S if there exists a $\delta > 0$ such that $B(x, \delta) \subset S$.

1.9. Definition. A set S is **open** if each of its points is an interior point of S.

1.10. Definition. The collection of all open subsets of E^n as defined above is called the usual **topology** for E^n. If S is a nonempty subset of E^n, then the **relative topology** on S is the collection of sets U such that $U = S \cap V$, where V is open in E^n.

It is easy to see that open balls, the whole space E^n, and the empty set $\varnothing$ are open sets. (See Exercise 1.5.) The union of any collection of open sets is an open set; the intersection of any finite collection of open sets is an open set.

1.11. Definition. A set S is **closed** if its complement $\sim S \equiv \mathsf{E}^n \sim S = \{x : x \in \mathsf{E}^n \text{ and } x \notin S\}$ is open.

It is easy to see that all finite sets of points in E^n, the whole space E^n, and the empty set $\varnothing$ are closed sets. (See Exercise 1.6.) The intersection of any collection of closed sets is a closed set; the union of any finite collection of closed sets is a closed set. It is possible for a set to be neither open nor closed.

1.12. Definition. A set S is **bounded** if there exists a $\delta > 0$ such that $S \subset B(\theta, \delta)$.

Examples. Consider the following subsets of E^2:

$$A = \{(x, y): (x-4)^2 + (y+2)^2 < 2\}$$
$$B = \{(x, y): 1 \leqslant x < 2 \text{ and } 1 < y \leqslant 3\}$$
$$C = \{(x, y): x + 2y \leqslant 4\}$$

(See Figure 1.3.) The sets A and B are both bounded. Set C is not bounded. Set A is the open ball with center $(3, -2)$ and radius $\sqrt{2}$, and is an open set. The set C is closed, and set B is neither open nor closed.

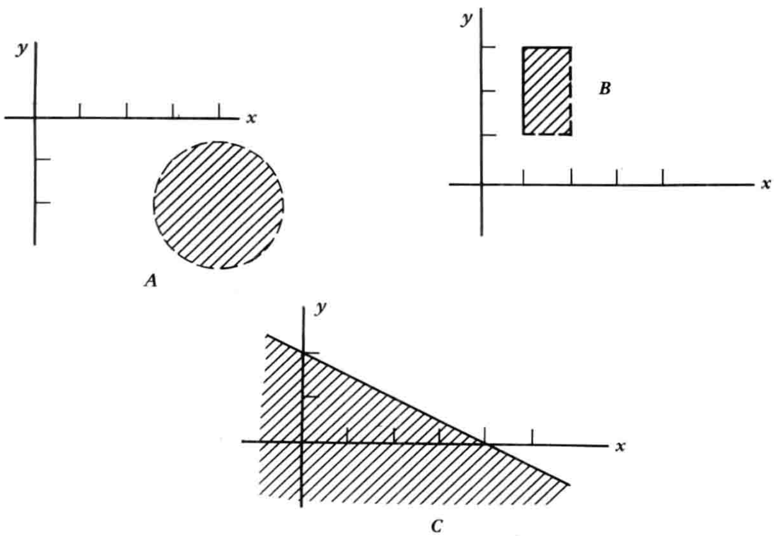

Figure 1.3.

1.13. Definition. The **interior** of a set S is the union of all the open sets contained in S. The **closure** of S is the intersection of all the closed sets containing S. The interior of S is denoted by int S and the closure of S by cl S.

It follows easily from the definitions that the interior of S is the set of all interior points of S. Also, a point x is in cl S iff for every $\delta > 0$, the open ball $B(x, \delta)$ contains at least one point of S.

1.14. Definition. A function f: $\mathsf{E}^n \to \mathsf{E}^m$ is **continuous** on E^n iff $f^{-1}(U)$ is an open subset of E^n whenever U is an open subset of E^m.

This definition of continuity is equivalent to the familiar ε-δ definition, which in terms of open balls becomes: f is continuous at $x \in \mathsf{E}^n$ iff for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $f(B(x, \delta)) \subset B(f(x), \varepsilon)$. If f is continuous at each point of a set A, then f is continuous on A. We also recall that if $\{x_k\}$ is a sequence of points in E^n which converges to x and if f is continuous on E^n, then $\{f(x_k)\}$ converges to $f(x)$.

The fundamental theorem relating the linear and topological structures is the following:

1.15. Theorem. Each of the following functions is continuous:

(a) f: $\mathsf{E}^n \times \mathsf{E}^n \to \mathsf{E}^n$ defined by $f(x, y) = x + y$.

(b) For any fixed $a \in \mathsf{E}^n$, f_a: $\mathsf{E}^n \to \mathsf{E}^n$ defined by $f_a(x) = a + x$.

(c) For any fixed $\lambda \in \mathsf{R}$, f_λ: $\mathsf{E}^n \to \mathsf{E}^n$ defined by $f_\lambda(x) = \lambda x$.

(d) For any fixed $x, y \in \mathsf{E}^n$, f: $\mathsf{R} \to \mathsf{E}^n$ defined by $f(\lambda) = \lambda x + (1 - \lambda)y$.

PROOF. (a) Given $\varepsilon > 0$, let $\delta = \varepsilon/2$. Let (x_0, y_0) be a point in $\mathsf{E}^n \times \mathsf{E}^n$. Then for any x and y in E^n, if $d((x, y), (x_0, y_0)) < \delta$ we have

$$d((x, y), (x_0, y_0)) = \left\{\left[d(x, x_0)\right]^2 + \left[d(y, y_0)\right]^2\right\}^{1/2}$$

so that $d(x, x_0) < \delta$ and $d(y, y_0) < \delta$. It follows that

$$\begin{aligned} d(x + y, x_0 + y_0) &\leqslant d(x + y, x + y_0) + d(x + y_0, x_0 + y_0) \\ &= d(y, y_0) + d(x, x_0) \\ &< \delta + \delta = \varepsilon. \end{aligned}$$

Thus if $(x, y) \in B((x_0, y_0), \delta)$, then $f(x, y) \in B(f(x_0, y_0), \varepsilon)$ and f is continuous at (x_0, y_0). Since (x_0, y_0) was an arbitrary point in E^n, f is continuous.

(b) This is just a special case of (a).

(c) Let $\varepsilon > 0$ and let $x \in \mathsf{E}^n$. If $\lambda \neq 0$, let $\delta = \varepsilon/|\lambda|$. Then for any $y \in \mathsf{E}^n$ such that $d(x, y) < \delta$ we have

$$d(f_\lambda(x), f_\lambda(y)) = d(\lambda x, \lambda y) = |\lambda|\, d(x, y) < |\lambda| \frac{\varepsilon}{|\lambda|} = \varepsilon.$$

If $\lambda = 0$, then $d(f_\lambda(x), f_\lambda(y)) = d(\theta, \theta) = 0 < \varepsilon$ for any δ. In both cases we have $f_\lambda(B(x, \delta)) \subset B(f_\lambda(x), \varepsilon)$, and f_λ is continuous.

(d) This proof is straightforward and is left as an exercise. ■

1.16. Definition. If $A, B \subset \mathsf{E}^n$ and $\lambda \in \mathsf{R}$, we define

$$A + B \equiv \{x + y: x \in A \text{ and } y \in B\}$$

$$\lambda A \equiv \{\lambda x: x \in A\}.$$

If A consists of a single point, $A \equiv \{x\}$, then we often write $x + B$ for $A + B$. The set $x + B$ is called a **translate** of B. The set λA is called a **scalar multiple** of A. If $\lambda \neq 0$, the set $x + \lambda A$ is said to be **homothetic** to A.

1.17. Theorem. Each set homothetic to an open set is open.

PROOF. For any $x \in \mathsf{E}^n$ and $\lambda \neq 0$, the function f given by $f(y) = x + \lambda y$ is continuous by Theorem 1.15. For $\lambda \neq 0$, its inverse $f^{-1}(z) = -(1/\lambda)x + (1/\lambda)z$ is also continuous. This implies that the original function maps open sets onto open sets. ■

1.18. Corollary. Each set homothetic to a closed set is closed.

PROOF. For $\lambda \neq 0$, the function $f(y) = x + \lambda y$ is one-to-one. Thus $f(\sim A) = \sim f(A)$ for each $A \subset \mathsf{E}^n$ and the corollary follows immediately from the definition of a closed set. ■

In trying to visualize the sum $A + B$ of two sets, it is often helpful to express the sum as a union of translates:

$$A + B = \bigcup_{x \in A} (x + B) = \bigcup_{y \in B} (A + y)$$

Examples
(See Figure 1.4.)

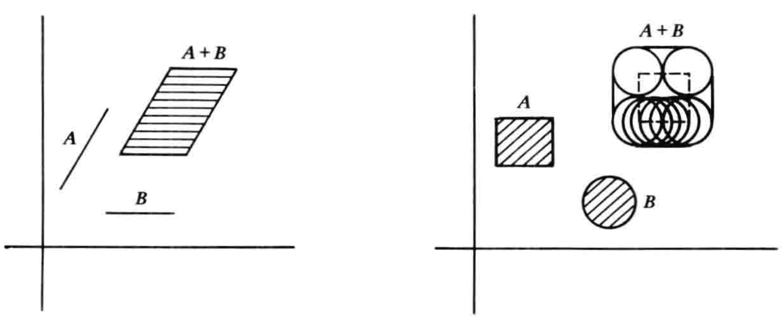

Figure 1.4.