Synopsis and Exercises for the Theory of Convex Sets
by G. D. Chakerian & J. R. Sangwine-Yager
April 28, 2009

Contents

List of Figures

Figure 1: The parallelogram law of addition in $\mathbf{R}^3$.

1 Introduction

In this section we introduce the fundamental properties of n-dimensional Euclidean space to be used throughout the course.

We shall denote the set of all real numbers by $\mathbf{R}$. If $n \geqslant 1$ is an integer, then n–*dimensional Euclidean space*, denoted by $\mathbf{R}^n$, consists of all ordered n-tuples of real numbers,

$$\mathbf{R}^n = \{(x_1, \ldots, x_n) \,:\, x_i \in \mathbf{R}\},$$

with some additional structure that will be described later. We shall call $(x_1, \ldots, x_n)$ a *point* of $\mathbf{R}^n$ and use the notation

$$\mathbf{x} = (x_1, \ldots, x_n).$$

The *origin* is given by $\mathbf{0} = (0, \ldots, 0)$.

Addition of points in $\mathbf{R}^n$ is defined as follows. If $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, then $\mathbf{x} + \mathbf{y}$ is defined by

$$\mathbf{x} + \mathbf{y} \stackrel{\text{def}}{=} (x_1 + y_1, \ldots, x_n + y_n).$$

In other words, each <u>coordinate</u> of $\mathbf{x} + \mathbf{y}$ is obtained by adding the corresponding coordinates of $\mathbf{x}$ and $\mathbf{y}$. As Figure 1 indicates, geometrically $\mathbf{x} + \mathbf{y}$ is the fourth vertex of the parallelogram with vertices at $\mathbf{0}$, $\mathbf{x}$, and $\mathbf{y}$.

Multiplication of a point in $\mathbf{R}^n$ by a scalar (real number) is defined as follows. If $\mathbf{x} = (x_1, \ldots, x_n)$ and $\lambda \in \mathbf{R}$, then $\lambda \mathbf{x}$ is defined by

$$\lambda \mathbf{x} \stackrel{\text{def}}{=} (\lambda x_1, \ldots, \lambda x_n).$$

Thus each coordinate of $\lambda \mathbf{x}$ is obtained by multiplying the corresponding coordinate of $\mathbf{x}$ by λ.

We see that $\mathbf{R}^n$ enjoys the properties of a vector space. In several instances we refer to the elements of $\mathbf{R}^n$ as *vectors*, especially when dealing with the properties involving linearity, such as linear independence. Usually we refer to the elements of $\mathbf{R}^n$ as <u>points</u>, in particular when dealing with the affine properties, such as affine independence, and especially in discussing metric properties such as distance between points. The linearity properties are dependent on the origin (the zero vector of the space) while the affine properties are independent of origin
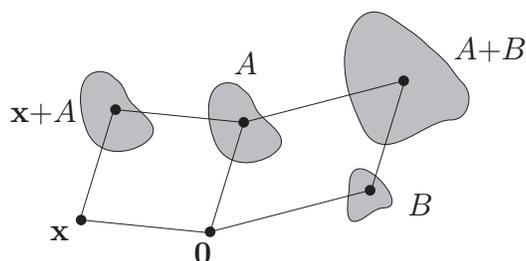
Figure 2: The Minkowski sum of sets and the translate of a set.

and invariant under the transformation known as translation. It may be helpful to keep in mind that each point of $\mathbf{R}^n$ has an associated *position vector* which we visualize intuitively as an arrow drawn from the origin to that point. We hope that the context will allow us to use the words "point" and "vector" interchangeably without confusing issues.

There is a natural way to define addition of subsets of $\mathbf{R}^n$. If A, B are any subsets of $\mathbf{R}^n$, the *Minkowski sum*, or *vector sum*, of A, B, is defined by

$$A + B \overset{\text{def}}{=} \{\mathbf{x} + \mathbf{y} : \mathbf{x} \in A \text{ and } \mathbf{y} \in B\}.$$

Note that a mechanical apparatus, based on the parallelogram law, could be constructed for obtaining $A + B$ from A and B (Figure 2). If $B = \{\mathbf{x}_0\}$, then $A + B = A + \{\mathbf{x}_0\}$ is a *translate of A* (Figure 2). We have a convention of writing $\mathbf{x}_0 + A$ for $\{\mathbf{x}_0\} + A$.

If $A \subseteq \mathbf{R}^n$ and $\lambda \in \mathbf{R}$, then λA is defined by

$$\lambda A \overset{\text{def}}{=} \{\lambda \mathbf{x} : \mathbf{x} \in A\}.$$

If $\lambda > 0$, this corresponds to "expanding" (or contracting!) A by the factor λ about the origin (Figure 3). In the special case $\lambda = -1$ we write

$$-A \overset{\text{def}}{=} (-1)A = \{-\mathbf{x} : \mathbf{x} \in A\},$$

and call $-A$ the *the reflection of A through the origin*, (Figure 3).

1.1   Exercises

1–1 In $\mathbf{R}^2$, let A be the square (including the interior and boundary points) with vertices at $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$ and let B be the open disk (not including boundary points) with center $(1,0)$ and radius 1. Sketch $A + B$.

1–2 The *difference set* of $A \subset \mathbf{R}^n$ is defined to be $A + (-A)$. By definition

$$A + (-A) = \{\mathbf{x} - \mathbf{y} : \mathbf{x}, \mathbf{y} \in A\} = \text{ the set of all "differences" of points of } A.$$

Show that $A + (-A)$ coincides with its reflection through $\mathbf{0}$. (A set S that coincides with its reflection through $\mathbf{0}$, i.e. such that $S = -S$, is said to be *centrally symmetric with center $\mathbf{0}$*.)
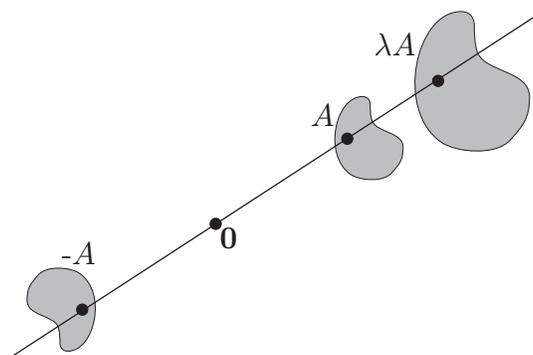
Figure 3: λA, the expansion of A where $\lambda > 1$, and $-A$, the reflection of A.

1–3 Let T be the (solid) equilateral triangle in $\mathbf{R}^2$ with vertices $(0,0)$, $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, $(-\frac{1}{2}, \frac{\sqrt{3}}{2})$. Sketch $T + (-T)$, and indeed show that $T + (-T)$ is a regular hexagon centered at the origin.

1–4 In $\mathbf{R}^2$, let A be the line segment $A = \{(t,t) \ : \ 0 \leqslant t \leqslant 1\}$ and B the line segment $B = \{(t,0) \ : \ 0 \leqslant t \leqslant 2\}$. Sketch $A + B$.

1–5 In $\mathbf{R}^2$, let D be the circular disk $\{(x,y) \ : \ x^2 + y^2 \leqslant 1\}$. Observe that if $A = \frac{1}{2}D$, then $A + (-A) = D$. Can you find $A \subseteq \mathbf{R}^2$, with $A \neq \frac{1}{2}D$, such that $A + (-A) = D$?

1–6 Prove that, $A + B = \bigcup_{\mathbf{x} \in A}(\mathbf{x} + B)$.

1–7 (a) Give an example of a set $A \subseteq \mathbf{R}$ such that $A + A = 2A$, but A is not an interval, a halfline, or all of $\mathbf{R}$.
   Remark. The line segment with endpoints $\mathbf{x}$, $\mathbf{y} \in \mathbf{R}^n$, for any n, is denoted by $\overline{\mathbf{xy}}$. A set $A \subset \mathbf{R}$ is said to be *midpoint convex* if whenever $\mathbf{x}$, $\mathbf{y} \in A$, then the midpoint of the segment $\overline{\mathbf{xy}}$ belongs to A.
   (b) Show that $A \subset \mathbf{R}$ is midpoint convex if an only if $A + A = 2A$.
   (c) Investigate, and try to characterize, those $A \subseteq \mathbf{R}$ such that $A + A = 2A$.

2   The Inner Product in $\mathbf{R}^n$

$\mathbf{R}^n$ is an example of an inner product space, that is, a vector space equipped with an inner product. We define the *inner product* of $\mathbf{x}$, $\mathbf{y}$ in $\mathbf{R}^n$ by

$$\langle \mathbf{x}, \mathbf{y} \rangle \overset{\text{def}}{=} \sum_{i=1}^{n} x_i y_i, \text{ where } \mathbf{x} = (x_1, \ldots, x_n) \text{ and } \mathbf{y} = (y_1, \ldots, y_n).$$

$\langle \mathbf{x}, \mathbf{y} \rangle$ is linear in each of the variables (bilinear). That is, if λ_1, $\lambda_2 \in \mathbf{R}$, then

$$\langle \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \mathbf{y} \rangle = \lambda_1 \langle \mathbf{x}_1, \mathbf{y} \rangle + \lambda_2 \langle \mathbf{x}_2, \mathbf{y} \rangle,$$

and similarly in the other variable.

The *norm* of $\mathbf{x} \in \mathbf{R}^n$ is

$$\|\mathbf{x}\| \stackrel{\text{def}}{=} \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}}.$$

Important properties of the norm:

> (i) $\|\mathbf{x}\| \geqslant 0$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$,
>
> (ii) $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$ for all $\lambda \in \mathbf{R}$, $\mathbf{x} \in \mathbf{R}^n$,
>
> (iii) $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$.

The proofs of properties (i) and (ii) are left for the reader in Exercise 2–1, below. Property (iii) requires for its proof the following:

Theorem 1 (The Cauchy-Schwarz Inequality) *If* $\mathbf{x}$, $\mathbf{y} \in \mathbf{R}^n$, *then*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant \|\mathbf{x}\| \, \|\mathbf{y}\|$$

with equality if and only if $\mathbf{x} = \lambda \mathbf{y}$ *or* $\mathbf{y} = \lambda \mathbf{x}$ *for some* $\lambda \in \mathbf{R}$.

Proof. Assume $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$. Then

$$\left\langle \mathbf{x} - \frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}\mathbf{y}, \mathbf{x} - \frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}\mathbf{y} \right\rangle \geqslant 0$$

which implies

$$\|\mathbf{x}\|^2 - 2\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}\langle \mathbf{x}, \mathbf{y} \rangle + \frac{\|\mathbf{x}\|^2}{\|\mathbf{y}\|^2}\|\mathbf{y}\|^2 \geqslant 0$$

which in turn implies

$$\langle \mathbf{x}, \mathbf{y} \rangle \leqslant \|\mathbf{x}\| \|\mathbf{y}\|.$$

To get $-\langle \mathbf{x}, \mathbf{y} \rangle \leqslant \|\mathbf{x}\| \|\mathbf{y}\|$, so $|\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant \|\mathbf{x}\| \, \|\mathbf{y}\|$, replace $\mathbf{x}$ with $-\mathbf{x}$.

The equality condition follows from properties (i) and (ii) of the norm.   ■

Proof of Property (iii), page 7.

$$
\begin{aligned}
\|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\
&= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\
&\leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2, \ \text{by the Cauchy-Schwarz inequality} \\
&= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad ■
\end{aligned}
$$

We have been investigating the properties of points in $\mathbf{R}^n$, but we may also consider $\mathbf{x}$ to be the position vector from the origin to the point $\mathbf{x}$. Two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^n$ are said
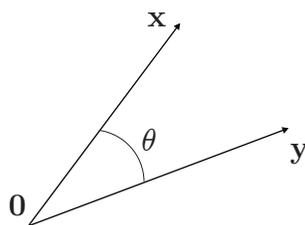
Figure 4: The angle between two vectors.

to be *independent* if neither is a scalar multiple of the other. Otherwise, they are said to be *dependent*. It follows that the vectors are dependent if they lie on the same line passing through the origin. If two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^n$ are independent, the Cauchy-Schwarz inequality implies

$$-1 < \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|} < 1.$$

Therefore there is a unique θ, $0 < \theta < \pi$, such that $\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$. We define this θ to be the *angle* between the vectors $\mathbf{x}$ and $\mathbf{y}$ (Figure 4). We then have

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos \theta.$$

$\mathbf{R}^n$ is also an example of a metric space. We define the *distance* $d(\mathbf{x}, \mathbf{y})$ between points $\mathbf{x}$ and $\mathbf{y}$ by

$$d(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}} = \|\mathbf{x} - \mathbf{y}\|.$$

Important properties of $d(\mathbf{x}, \mathbf{y})$:

> (i) $d(\mathbf{x}, \mathbf{y}) \geqslant 0$, and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$,
>
> (ii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$,
>
> (iii) $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geqslant d(\mathbf{x}, \mathbf{z})$, *The Triangle Inequality* .

These properties follow readily from the properties of the norm.

2.1   Exercises

In the following, Exercises 2–2 to 2–6 will guide you through a proof of Hölder's inequality (which generalizes the Cauchy-Schwarz inequality) and then a proof of an important inequality of Minkowski.

2–1 Prove properties (i) and (ii) of the norm, page 7.

2–2 (a) Suppose $0 < \alpha < 1$ and $t \geqslant 0$. Prove that $t^\alpha - \alpha t \leqslant 1 - \alpha$, with equality if and only if $t = 1$.

(b) Suppose $\alpha > 1$ and $t \geqslant 0$. Prove that $t^\alpha - \alpha t \geqslant 1 - \alpha$, with equality if and only if $t = 1$.

[Hint: Study the graph of $f(t) = t^\alpha - \alpha t$.]

2–3 Suppose a, $b > 0$, and p, $q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$. Prove that

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q.$$

[Hint: Let $t = \frac{a^p}{b^q}$ and $\alpha = \frac{1}{p}$ in Exercise 2–2 (a).]

2–4 Suppose $a_1, \ldots, a_n \geqslant 0$, $b_1, \ldots, b_n \geqslant 0$, with $\sum_{i=1}^n a_i^p = \sum_{i=1}^n b_i^q = 1$, where p, $q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$. Prove that

$$\sum_{i=1}^n a_i b_i \leqslant 1.$$

2–5 Suppose $x_1, \ldots, x_n \geqslant 0$, $y_1, \ldots, y_n \geqslant 0$, p, $q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$. Prove *Hölder's inequality*:

$$\boxed{\sum_{i=1}^n x_i y_i \leqslant \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}}.}$$

[Hint: Let $a_i = x_i / (\sum x_i^p)^{\frac{1}{p}}$, $b_i = y_i / (\sum y_i^q)^{\frac{1}{q}}$]

Why is this a generalization of the Cauchy-Schwarz inequality?

2–6 Suppose $x_1, \ldots, x_n$, $y_1, \ldots, y_n \in \mathbf{R}$ and $p \geqslant 1$. Prove *Minkowski's inequality*:

$$\boxed{\left(\sum_{i=1}^n |x_i + y_i|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p\right)^{\frac{1}{p}}.}$$

[Use the idea of F. Riesz, see Hardy et al. (1952, p. 24). First show that it suffices to consider the case where $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ are nonnegative. Then write

$$\sum_{i=1}^n (x_i + y_i)^p = \sum_{i=1}^n x_i(x_i + y_i)^{p-1} + \sum_{i=1}^n y_i(x_i + y_i)^{p-1},$$

and apply Hölder's inequality to the terms on the righthand side, with $q = \frac{p}{p-1}$.]

2–7 The inequality in Exercise 2–2 (b) can be used to prove the famous *inequality between the arithmetic and geometric means* of n positive real numbers. If $a_1, \ldots, a_n > 0$, this

inequality asserts that the arithmetic mean is greater than or equal to the geometric mean. That is,

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geqslant (a_1 a_2 \cdots a_n)^{\frac{1}{n}}.$$

with equality if and only if $a_1 = a_2 = \cdots = a_n$.

The following idea for a proof is from Akerberg (1963):

Rewrite 2–2 (b) in the form $t(\alpha - t^{\alpha-1}) \leqslant \alpha - 1$. Let $A = \frac{a_1 + a_2 + \cdots + a_n}{n}$, and substitute $t = (a_1/A)^{\frac{1}{n}}$ and $\alpha = n$. Show that this gives

$$\left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^n \geqslant a_1 \left(\frac{a_1 + a_2 + \cdots + a_n}{n-1}\right)^{n-1}.$$

But a repetition of this, applied to $a_2, \ldots, a_n$ on the righthand side, gives then

$$\left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^n \geqslant a_1 a_2 \left(\frac{a_1 + a_2 + \cdots + a_n}{n-2}\right)^{n-2}.$$

Continuing, we have

$$\left(\frac{a_1 + a_2 + \cdots + a_n}{n}\right)^n \geqslant a_1 a_2 \cdots a_n.$$

2–8 (a) Suppose $f(t)$ is continuous and strictly increasing for $t \geqslant 0$, and $f(0) = 0$. Let g be the inverse function of f. If a, $b > 0$, prove *Young's inequality*:

$$ab \leqslant \int_0^a f(x)\,dx + \int_0^b g(y)\,dy.$$

[Hint: Interpret the various quantities as areas. Sketch the graph of f.]

(b) Recall that for a, $b > 0$, p, $q > 0$ with $\frac{1}{p} + \frac{1}{q} = 1$, the inequality

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{q}b^q,$$

proved in Exercise 2–3, was crucial for the proof of Hölder's inequality. Derive this inequality from Young's inequality.

3   Combinations

3.1   Linear, Affine and Convex Combinations

If $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$, then

- a *linear combination* of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is

$$\lambda_1 \mathbf{x}_1 + \cdots \lambda_k \mathbf{x}_k, \text{ for } \lambda_1, \ldots, \lambda_k \in \mathbf{R}.$$

- an *affine combination* of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is

$$\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k, \text{ where } \lambda_1 + \cdots + \lambda_k = 1.$$

- a *convex combination* of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is

$$\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k, \text{ where } \lambda_1 + \cdots + \lambda_k = 1 \text{ and } \lambda_1 \geqslant 0, \ldots, \lambda_k \geqslant 0.$$

Thus a convex combination is an affine combination with nonnegative coefficients, and an affine combination is a linear combination whose coefficients sum to 1.

Previously, page 8, we defined linear independence for a pair of vectors. More generally, the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$ are said to be *linearly independent* if no one of them is a linear combination of the others. Equivalently, $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$ are linearly independent if

$$\sum_{i=1}^{k} \lambda_i \mathbf{x}_i = \mathbf{0} \text{ for } \lambda_1, \ldots, \lambda_k \in \mathbf{R} \text{ implies } \lambda_1 = \cdots = \lambda_k = 0.$$

Otherwise, they are said to be *linearly dependent.* The points $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$ are said to be *affinely independent* if no one of them is a affine combination of the others. An equivalent formulation is that $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$ are affinely independent if

$$\sum_{i=1}^{k} \lambda_i \mathbf{x}_i = \mathbf{0} \text{ for } \lambda_1, \ldots, \lambda_k \in \mathbf{R} \text{ and } \sum_{i=1}^{k} \lambda_i = 0 \text{ implies } \lambda_1 = \cdots = \lambda_k = 0.$$

Otherwise, they are said to be *affinely dependent.*

In the following examples, suppose $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{R}^n$.

Example 3–A. The sum, $\mathbf{x}_1 + \mathbf{x}_2$, is a linear combination.

Example 3–B. The *midpoint* of the line segment with endpoints $\mathbf{x}_1$ and $\mathbf{x}_2$ is $\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$. Thus the midpoint is a convex combination.

A "Physical" Interpretation. Place masses m_i at $\mathbf{x}_i \in \mathbf{R}^n$ for $i = 1, \ldots, k$. Then

$$\text{the center of mass } \overset{\text{def}}{=} \frac{m_1 \mathbf{x}_1 + \cdots + m_k \mathbf{x}_k}{m_1 + \cdots + m_k}.$$

Note this is a convex combination of $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

3.2 Centroids

If $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R^n}$ and the masses in the previous physical interpretation are all equal, we obtain the

$$\text{centroid of } \mathbf{x}_1, \ldots, \mathbf{x}_k \overset{\text{def}}{=} \frac{1}{k}(\mathbf{x}_1 + \cdots + \mathbf{x}_k).$$

Figure 5 illustrates the following examples.
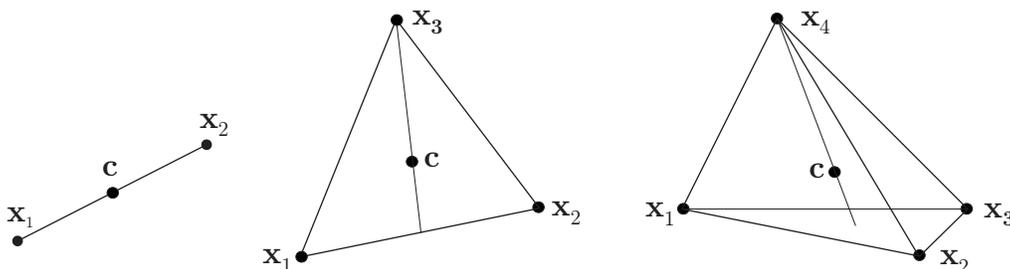
Figure 5: **c** is the centroid of 2, 3 and 4 points respectively.

Example 3–C. The centroid of $\mathbf{x}_1$, $\mathbf{x}_2 \in \mathbf{R}^n$ is the midpoint of the line joining them.

Example 3–D. Three distinct, non-collinear points $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3 \in \mathbf{R}^n$ determine a plane. Their centroid,

$$\frac{1}{3}(\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3) = \mathbf{x}_3 + \frac{2}{3}\left(\frac{\mathbf{x}_1 + \mathbf{x}_2}{2} - \mathbf{x}_3\right)$$

is the familiar centroid of the triangle with vertices $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. Here, the centroid of the point is intersection of the medians of the triangle and these medians trisect one another.

Example 3–E. Four affinely independent points $\mathbf{x}_1$, $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4 \in \mathbf{R}^3$ determine a *3–dimensional simplex*. (For more on simplices see Exercise 8–10, page 30.) Their centroid is

$$\frac{1}{4}(\mathbf{x}_1 + \cdots + \mathbf{x}_4) = \mathbf{x}_4 + \frac{3}{4}\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_3}{3} - \mathbf{x}_4\right).$$

One should avoid confusion between the centroid of the vertices and the centroid of the solid, which happen to coincide for simplices. The centroid of a solid is defined in terms of certain integrals, but we shall not go into this here. An excellent introduction may be found in Bonnesen & Fenchel (1934, §2).

3.3 Exercises

3–1 An *n–dimensional simplex* is determined by $n + 1$ affinely independent points $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_{n+1} \in \mathbf{R}^n$. (For more on simplices see Exercise 8–10, page 30.) Let **c** be the centroid of $\mathbf{x}_1, \ldots, \mathbf{x}_{n+1}$ and let $\mathbf{c}_1$ be the centroid of $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Prove that **c** is the following convex combination of $\mathbf{c}_1$ and $\mathbf{x}_{n+1}$:

$$\mathbf{c} = \frac{n}{n+1}\mathbf{c}_1 + \frac{1}{n+1}\mathbf{x}_{n+1}.$$

Remark. This tells us that the centroid of an *n*–dimensional simplex divides the line segment joining any vertex to the centroid of the $(n - 1)$–dimensional simplex determined by the other n points in the ratio $n : 1$. The next exercise can be viewed as a generalization of this result.
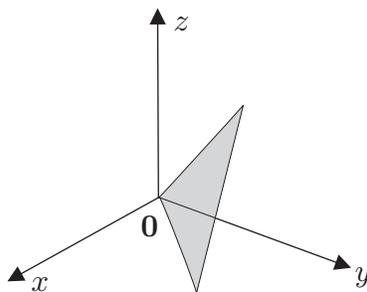
Figure 6: The plane $x - y + z = 0$ in $\mathbf{R}^3$, in Example 4–A.

3–2 Suppose $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ and $1 < r < k$ for some integer r. Let $\mathbf{c}$ be the centroid of $\mathbf{x}_1, \ldots, \mathbf{x}_k$, let $\mathbf{c}_1$ be the centroid of $\mathbf{x}_1, \ldots, \mathbf{x}_r$, and let $\mathbf{c}_2$ be the centroid of $\mathbf{x}_{r+1}, \ldots, \mathbf{x}_k$. Prove that
$$\mathbf{c} = \frac{r}{k}\mathbf{c}_1 + (1 - \frac{r}{k})\mathbf{c}_2.$$

3–3 Use Exercise 3–1 to deduce that the medians of a triangle intersect at one point.

3–4 Show that the points $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ are affinely independent if and only if the vectors $\mathbf{x}_2 - \mathbf{x}_1, \ldots, \mathbf{x}_k - \mathbf{x}_1$ are linearly independent.

3–5 In the definition of an <u>affine combination</u> we have $\sum_{i=1}^{k} \lambda_i = 1$, but in the definition of <u>affine independence</u> we have $\sum_{i=1}^{k} \lambda_i = 0$. Explain why the latter set of λ_i's summing to zero is not inconsistent with the former set summing to one.

4   Linear and Affine Hulls

The *linear hull* (or *linear span*) of $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ consists of the set of all possible linear combinations of $\mathbf{x}_1, \ldots, \mathbf{x}_k$:

$$\text{linear hull of } \mathbf{x}_1, \ldots, \mathbf{x}_k \overset{\text{def}}{=} \{\lambda_1 \mathbf{x}_1 + \cdots \lambda_k \mathbf{x}_k \ : \ \lambda_1, \ldots, \lambda_k \in \mathbf{R}\}.$$

We say that $S \subseteq \mathbf{R}^n$, $S \neq \emptyset$, is a *subspace* if and only if $\mathbf{x}_1, \mathbf{x}_2 \in S$ implies $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in S$ for all $\lambda_1, \lambda_2 \in \mathbf{R}$. That is, S is "closed" under taking linear combinations of pairs of points of S. The *dimension of a subspace* is the number of linearly independent vectors required to span the subspace.

Example 4–A. In $\mathbf{R}^3$ let $\mathbf{x}_1 = (1, 1, 0)$, $\mathbf{x}_2 = (0, 1, 1)$. Then the linear hull of $\mathbf{x}_1$, $\mathbf{x}_2$ is

$$\{\lambda \mathbf{x}_1 + \lambda_2 \mathbf{x}_2\} = \{(\lambda_1, \lambda_1 + \lambda_2, \lambda_2) \ : \ \lambda_1, \lambda_2 \in \mathbf{R}\}.$$

Note that this is a plane through $\mathbf{0}$ with equation $x - y + z = 0$ (Figure 6). It is a 2-dimensional subspace of $\mathbf{R}^3$.

The example illustrates a general property:

Theorem 2 *If* $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$, *then their linear hull is*

 (i) *a subspace of* $\mathbf{R}^n$;

 (ii) *in fact, the "smallest" subspace of* $\mathbf{R}^n$ *containing* $\mathbf{x}_1, \ldots, \mathbf{x}_k$. *That is, if S is a subspace of* $\mathbf{R}^n$ *containing* $\mathbf{x}_1, \ldots, \mathbf{x}_k$, *then S contains the linear hull of* $\mathbf{x}_1, \ldots, \mathbf{x}_k$;

 (iii) *the intersection of all subspaces of* $\mathbf{R}^n$ *which contain* $\mathbf{x}_1, \ldots, \mathbf{x}_k$. *That is,*

$$\text{linear hull of } \mathbf{x}_1, \ldots, \mathbf{x}_k = \bigcap \{ S \ : \ \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subseteq S \text{ and } S \text{ is a subspace of } \mathbf{R}^n \}.$$

The proof is left to the reader.

The *affine hull* of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ consists of all affine combinations of $\mathbf{x}_1, \ldots, \mathbf{x}_k$:

$$\text{affine hull of } \mathbf{x}_1, \ldots, \mathbf{x}_k \overset{\text{def}}{=} \{ \lambda_1 \mathbf{x}_1 + \cdots \lambda_k \mathbf{x}_k \ : \ \lambda_1 + \cdots + \lambda_k = 1 \}.$$

A *flat* in $\mathbf{R}^n$ is any translate of a subspace in $\mathbf{R}^n$. Thus if F is a flat, then there exists a subspace S and $\mathbf{x}_0$ such that

$$F = \mathbf{x}_0 + S.$$

You may think of a flat as a subset closed under taking the straight line through each pair of its points. The *dimension of a flat* is the dimension of the subspace of which it is a translate.

Example 4–B. The affine hull of two distinct points is the straight line through those two points. Thus it is a one-dimensional flat. Note that we have two descriptions of the straight line through $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{R}^n$:

$$\{ \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \ : \ \lambda_1 + \lambda_2 = 1 \}$$

or

$$\{ (1 - \lambda)\mathbf{x}_1 + \lambda \mathbf{x}_2 = \mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1) \ : \ -\infty < \lambda < \infty \}.$$

Example 4–C. In $\mathbf{R}^3$ let $\mathbf{x}_1 = (1, 0, 0)$, $\mathbf{x}_2 = (0, 1, 0)$, $\mathbf{x}_3 = (0, 0, 1)$. Then

$$
\begin{aligned}
\text{affine hull of } \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \ &= \ \{ \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \lambda_3 \mathbf{x}_3 \ : \ \lambda_1 + \lambda_2 + \lambda_3 = 1 \} \\
&= \ \{ (\lambda_1, \lambda_2, \lambda_3) \ : \ \lambda_1 + \lambda_2 + \lambda_3 = 1 \} \\
&= \ \text{the plane with equation } x + y + z = 1.
\end{aligned}
$$

The plane $x + y + z = 1$ is a 2–dimensional flat in $\mathbf{R}^3$ (Figure 7).

These examples illustrate the general property:

Theorem 3 *If* $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$, *then their affine hull is*

 (i) *a flat in* $\mathbf{R}^n$;
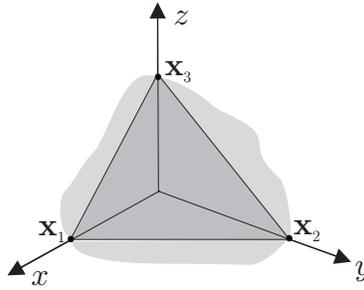
Figure 7: The plane $x + y + z = 1$ in $\mathbf{R}^3$, in Example 4–C.

(ii) in fact, the "smallest" flat in $\mathbf{R}^n$ containing $\mathbf{x}_1, \ldots, \mathbf{x}_k$. That is, if F is a flat in $\mathbf{R}^n$ containing $\mathbf{x}_1, \ldots, \mathbf{x}_k$, then F contains the affine hull of $\mathbf{x}_1, \ldots, \mathbf{x}_k$;

(iii) the intersection of all flats in $\mathbf{R}^n$ which contain $\mathbf{x}_1, \ldots, \mathbf{x}_k$. That is,

$$\text{affine hull of } \mathbf{x}_1, \ldots, \mathbf{x}_k = \bigcap \{F : \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subseteq F \text{ and } F \text{ is a flat of } \mathbf{R}^n\}.$$

Proof. For the proof of (i), let

$$
\begin{aligned}
F &= \{\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k : \lambda_1 + \cdots + \lambda_k = 1\} \\
&= \{(1 - \lambda_2 - \cdots - \lambda_k)\mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_k \mathbf{x}_k : \lambda_2, \ldots, \lambda_k \in R\} \\
&= \{\mathbf{x}_1 + \lambda_2(\mathbf{x}_2 - \mathbf{x}_1) + \cdots + \lambda_k(\mathbf{x}_k - \mathbf{x}_1) : \lambda_2, \ldots, \lambda_k \in R\} \\
&= \mathbf{x}_1 + S,
\end{aligned}
$$

where S is the subspace $\{\lambda_2(\mathbf{x}_2 - \mathbf{x}_1) + \cdots + \lambda_k(\mathbf{x}_k - \mathbf{x}_1) : \lambda_2, \ldots, \lambda_k \in R\}$.

For the proof of (ii), let $F = \mathbf{x}_0 + S$, for some subspace S, be any flat such that $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subset F$. Since $\mathbf{x}_i \in F$, we have $\mathbf{x}_i = \mathbf{x}_0 + \mathbf{y}_i$ for some $\mathbf{y}_i \in S$, $i = 1, \ldots, k$. Note then that any linear combination of $\mathbf{y}_1, \ldots, \mathbf{y}_k$ belongs to S. If $\lambda_1 + \cdots + \lambda_k = 1$, then

$$
\begin{aligned}
\sum_{i=1}^{k} \lambda_i \mathbf{x}_i &= \sum_{i=1}^{k} \lambda_i(\mathbf{x}_0 + \mathbf{y}_i) \\
&= \left(\sum_{i=1}^{k} \lambda_i \mathbf{x}_0\right) + \left(\sum_{i=1}^{k} \lambda_i \mathbf{y}_i\right) \\
&= \left(\sum_{i=1}^{k} \lambda_i\right)\mathbf{x}_0 + \left(\sum_{i=1}^{k} \lambda_i \mathbf{y}_i\right) \\
&= \mathbf{x}_0 + \left(\sum_{i=1}^{k} \lambda_i \mathbf{y}_i\right).
\end{aligned}
$$

The latter sum is a linear combination of $\mathbf{y}_1, \ldots, \mathbf{y}_k$ and hence an element of S. This shows that every affine combination of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is in F; thus the affine hull is a subset of F. This completes the proof of (ii).

Part (iii) follows from (i) and (ii).   ■
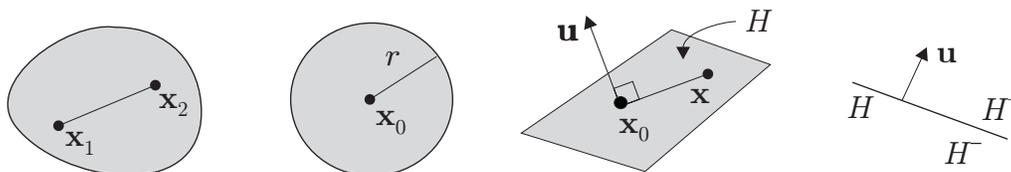
Figure 8: A convex set, and examples of a ball, hyperplane and halfspaces.

4.1   Exercises

4–1 (a) Let $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ with $k \geqslant n+2$. It is known that any $n+1$ or more vectors in $\mathbf{R}^n$ are linearly independent. Deduce that $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are affinely independent (Exercise 3–4, page 13).
(b) Show that $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ are affinely independent if and only if one of them is contained in the affine hull of the others.

4–2 In $\mathbf{R}^3$, let $\mathbf{x}_1 = (1,0,0)$, $\mathbf{x}_2 = (-1,0,0)$, $\mathbf{x}_3 = (0,1,0)$, $\mathbf{x}_4 = (0,-1,0)$.
(a) Describe geometrically the linear hull of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$.
(b) Describe geometrically the affine hull of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$.

5   Convex Sets

Having used linear combinations to define the linear hull and affine combinations to define the affine hull, we now expect to use convex combinations to define the convex hull. We also expect to state a theorem which tells us that the convex hull of a set of points is the "smallest object" that contains the set. This section identifies that object.

$K \subset \mathbf{R}^n$ is *convex* if and only if whenever $\mathbf{x}_1$, $\mathbf{x}_2 \in K$, then all the points of the line segment with endpoints $\mathbf{x}_1$, $\mathbf{x}_2$ belong to K (Figure 8). In other words, K is convex if and only if whenever $\mathbf{x}_1$, $\mathbf{x}_2 \in K$, then $(1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2 \in K$ for all $0 \leqslant \lambda \leqslant 1$. The *dimension of a convex set* is the dimension of its affine hull.

Example 5–A. $\mathbf{R}^n$ is an n–dimensional convex set.

Example 5–B. Any flat in $\mathbf{R}^n$ is convex.

Example 5–C. Let $\mathbf{x}_0 \in \mathbf{R}^n$ and $r > 0$. The *closed ball* of radius r centered at $\mathbf{x}_0$

$$B(\mathbf{x}_0, r) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) \leqslant r\}$$

is an n–dimensional convex set (Figure 8).

Example 5–D. The *open ball*
$$\{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) < r\}$$

is also an n–dimensional convex set.

Example 5–E. Let $\mathbf{x}_0 \in \mathbf{R}^n$ and $\|\mathbf{u}\| = 1$ ($\mathbf{u}$ is a "direction"). Then the set

$$H = \{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle = 0\}$$

is a *hyperplane* passing through $\mathbf{x}_0$ and having unit normal $\mathbf{u}$ (Figure 8). Note that a hyperplane is an $(n-1)$–dimensional flat.

$$H^+ = \{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle \geqslant 0\} \text{ and}$$

$$H^- = \{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle \leqslant 0\}$$

are the two *closed halfspaces* defined by H (Figure 8). Any closed halfspace is an n–dimensional convex set. The *open halfspaces* defined by H are

$$\{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle > 0\} \text{ and}$$

$$\{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle < 0\}.$$

These are also an n–dimensional convex sets.

5.1   Exercises

5–1 Prove that every halfspace is convex. [Hint: Suppose H^+ is a closed halfspace given by $\{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle \geqslant 0\}$. Given $\mathbf{x}_1, \mathbf{x}_2 \in H^+$, it is required to show that $(1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2 \in H^+$ when $0 \leqslant \lambda \leqslant 1$.

5–2 Prove that every ball is convex. [Hint: Suppose $B = \{\mathbf{x} \in \mathbf{R}^n : d(\mathbf{x}, \mathbf{x}_0) \leqslant r\}$. Thus if $\mathbf{x}_1, \mathbf{x}_2 \in B$, then $\|\mathbf{x}_1 - \mathbf{x}_0\| \leqslant r$ and $\|\mathbf{x}_2 - \mathbf{x}_0\| \leqslant r$. One needs to show in that case $(1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2 \in B$ when $0 \leqslant \lambda \leqslant 1$. A crucial observation that should help you is

$$((1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2) - \mathbf{x}_0 = (1-\lambda)(\mathbf{x}_1 - \mathbf{x}_0) + \lambda(\mathbf{x}_2 - \mathbf{x}_0).]$$

5–3 A very useful tool in establishing that certain sets are convex is the following fact: *"The intersection of convex sets is convex."* Prove that for any collection of convex sets $\{K_\alpha : \alpha \in \mathcal{A}\}$, the set $\bigcap_{\alpha \in \mathcal{A}} K_\alpha$ is again a convex set.

5–4 In $\mathbf{R}^n$, let I^n be the *n–dimensional cube* defined by

$$I^n \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbf{R}^n : -1 \leqslant x_i \leqslant 1, \ i = 1, \ldots, n\}.$$

(a) Sketch this set in case $n = 1, 2, 3$.
(b) I^n is the intersection of certain closed halfspaces. What are they?
(c) Why does part (b) show that I^n is convex?

5–5 In $\mathbf{R}^n$, the *n–dimensional crosspolytope* is defined by

$$C_n \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbf{R}^n \ : \ |x_1| + \cdots + |x_n| \leqslant 1\}.$$

    (a) Sketch this set in case $n = 1, 2, 3$.
    (b) Prove that the n–dimensional crosspolytope is convex.

5–6 Let $a_1, \ldots, a_n > 0$. Prove that the (solid) *n–dimensional ellipsoid*

$$E = \{\mathbf{x} \in \mathbf{R}^n \ : \ \sum_{i=1}^n \frac{x_i^2}{a_i^2} \leqslant 1\}$$

is convex. [Hint: Proceed directly. Suppose $\mathbf{x}, \mathbf{y} \in E$. Then $\sum \frac{x_i^2}{a_i^2} \leqslant 1$ and $\sum \frac{y_i^2}{a_i^2} \leqslant 1$. If $0 \leqslant \lambda \leqslant 1$ then one needs to show that $(1 - \lambda)\mathbf{x} + \lambda \mathbf{y} \in E$, that is,

$$\sum_{i=1}^n \frac{((1 - \lambda)x_i + \lambda y_i)^2}{a_i^2} \leqslant 1.$$

Apply Minkowski's inequality, Exercise 2–6, to the square root of the sum.]

5–7 Describe all the convex subsets of $\mathbf{R}$.

5–8 (a) Prove that if K_1 and K_2 are convex, then $K_1 + K_2$ is convex.
    (b) Prove that if $K_1, \ldots, K_r$ are convex, then $K_1 + \cdots + K_r$ is convex.

5–9 Use Exercise 5–8 to show in another way that the n–dimensional cube I^n described in Exercise 5–4 is convex [Hint: $I^n = I_1 + \cdots + I_n$, where $I_i = ?$]

5–10 Prove that if K is convex, then $K + K = 2K$.

6   Convex Hulls and Polytopes

6.1   The Convex Hull

Rather than give the definition of the convex hull of a finite set of points, we present a more general definition.

    The *convex hull* of $A \subset \mathbf{R}^n$, denoted $\operatorname{conv}(A)$, is defined to be the set of all convex combinations of finitely many elements of A. That is,

$$\operatorname{conv}(A) \stackrel{\text{def}}{=} \{\lambda_1 \mathbf{x}_1 + \cdots \lambda_k \mathbf{x}_k \ :$$

$$k \text{ a positive integer}, \ \mathbf{x}_1, \ldots, \mathbf{x}_k \in A, \ \sum_{i=1}^k \lambda_i = 1, \ \text{and } \lambda_i \geqslant 0, i = 1, \ldots, k\}.$$
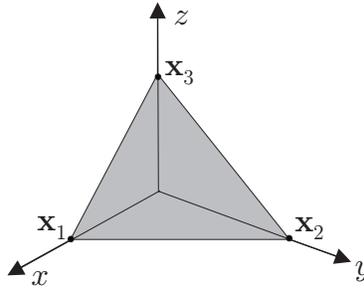
Figure 9: The set of convex combinations of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, in Example 6–B.

If A is a finite set, k can be taken to be the number of elements of A. However, if A is infinite, k is arbitrary. Analogous to our earlier definitions of hulls, we have

$$\mathrm{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \{\lambda_1 \mathbf{x}_1 + \cdots \lambda_k \mathbf{x}_k \; : \; \sum_{i=1}^{k} \lambda_i = 1, \; \text{and } \lambda_i \geqslant 0, i = 1, \ldots, k\}.$$

When using set notation, we have a convention of writing $\mathrm{conv}\{\ \}$ for $\mathrm{conv}(\{\ \})$.

Example 6–A. The line segment with endpoints $\mathbf{x}_1$, $\mathbf{x}_2 \in \mathbf{R}^n$ consists of the set of all convex combinations of $\mathbf{x}_1$ and $\mathbf{x}_2$. That is,

$$\{(1 - \lambda)\mathbf{x}_1 + \lambda \mathbf{x}_2 \; : \; 0 \leqslant \lambda \leqslant 1\}.$$

Example 6–B. In $\mathbf{R}^3$ let $\mathbf{x}_1 = (1, 0, 0)$, $\mathbf{x}_2 = (0, 1, 0)$, $\mathbf{x}_3 = (0, 0, 1)$. Then the convex combinations fill out the closed triangle with the vertices $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ (Figure 9).

We are now ready for theorem which we anticipated from Section 4.

Theorem 4 *If $A \subset \mathbf{R}^n$, then its convex hull is*

(i) a convex set in $\mathbf{R}^n$;

(ii) in fact, the "smallest" convex set in $\mathbf{R}^n$ containing A. That is, if K is a convex set in $\mathbf{R}^n$ containing A, then K contains the convex hull of A;

(iii) the intersection of all convex sets in $\mathbf{R}^n$ which contain A. That is,

$$\mathrm{conv}(A) = \bigcap\{K \; : \; A \subseteq K \text{ and } K \text{ is convex set in } \mathbf{R}^n\}.$$

Proof. For the proof of (i), we must show that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are convex combinations of points in A, then so is $(1 - \lambda)\mathbf{x}_1 + \lambda \mathbf{x}_2$ for $0 \leqslant \lambda \leqslant 1$. Showing this is a nice algebraic exercise for the reader (Exercise 6–1).

To prove (ii), assume that K is a convex set in $\mathbf{R}^n$ and $A \subset K$. We will proceed by induction to show that if $\mathbf{x}_1, \ldots, \mathbf{x}_k \in A$, then $\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k \in K$ for all $\lambda_1, \ldots, \lambda_k$ with

$\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geqslant 0$, $i = 1, \dots, k$. The statement is clearly true if $k = 1$. (Notice, by the convexity of K, it is also true for $k = 2$.) Assume it is true for $k - 1$. Then, assuming $\lambda_1 \neq 1$,

$$\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k \;=\; \lambda_1 \mathbf{x}_1 + (1 - \lambda_1)\Big(\frac{\lambda_2}{1 - \lambda_1}\mathbf{x}_2 + \cdots + \frac{\lambda_k}{1 - \lambda_1}\mathbf{x}_k\Big)$$

$$= \lambda_1 \mathbf{x}_1 + (1 - \lambda_1)\mathbf{y}, \quad \text{where } \mathbf{y} = \frac{\lambda_2}{1 - \lambda_1}\mathbf{x}_2 + \cdots + \frac{\lambda_k}{1 - \lambda_1}\mathbf{x}_k.$$

Now

$$\frac{\lambda_2}{1 - \lambda_1} + \cdots + \frac{\lambda_k}{1 - \lambda_1} = \frac{1}{1 - \lambda_1}(\lambda_2 + \cdots + \lambda_k) = 1,$$

so $\mathbf{y}$ is a convex combination of $k - 1$ elements of A. By our induction hypothesis, it is an element of K. But $\lambda_1 \mathbf{x}_1 + (1 - \lambda_1)\mathbf{y}$ has now been shown to be a convex combination of two elements of K and is also in K. This completes the proof of (ii).

As before, (iii) follows from (i) and (ii). ∎

6.2 Hulls Revisited

We now have the following summary.

If $A \subset \mathbf{R}^n$, then

$$\left.\begin{array}{c} \text{linear hull} \\ \text{affine hull} \\ \text{convex hull} \end{array}\right\} \text{ of } A \stackrel{\text{def}}{=} \text{ the set of all } \left\{\begin{array}{c} \text{linear} \\ \text{affine} \\ \text{convex} \end{array}\right\} \text{ combinations of finite subsets of } A,$$

and

$$\left.\begin{array}{c} \text{linear hull} \\ \text{affine hull} \\ \text{convex hull} \end{array}\right\} \text{ of } A \stackrel{\text{Theorem}}{=} \text{ the intersection of all } \left\{\begin{array}{c} \text{subspaces} \\ \text{flats} \\ \text{convex sets} \end{array}\right\} \text{ containing } A.$$

6.3 Convex Polytopes

A *convex polytope* is the convex hull of a finite set of points. In $\mathbf{R}^2$, a convex polytope is called a *convex polygon*.

Example 6–C. In $\mathbf{R}^2$ the convex polytopes are points, line segments, and convex polygons.

Example 6–D. The cube in $\mathbf{R}^3$ is a familiar example (Exercise 5–4). The 3-dimensional crosspolytope

$$C_3 = \{(x, y, z) \;:\; |x| + |y| + |z| \leqslant 1\}$$

is another example (Exercise 5–5). Note that C_3 is the familiar "regular octahedron".

Theorem 5 *The n–dimensional crosspolytope (Exercise 5–5)*

$$C_n = \{\mathbf{x} \in \mathbf{R}^n \ : \ |x_1| + \cdots + |x_n| \leqslant 1\}$$

is the convex hull of the $2n$ points $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$, where $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, etc.

Proof. Proceed by induction as follows. Clearly the result is true for $n = 1$. Supposing it true for dimension $n - 1$, we show that it is true for dimension n. If $\mathbf{x} \in C_n$, consider first the case where $\mathbf{x} = (x_1, \ldots, x_n)$ satisfies $0 \leqslant x_n \leqslant 1$. If $x_n = 1$, then $\mathbf{x} = \mathbf{e}_n$ and there is nothing to prove. If $0 \leqslant x_n < 1$, let

$$\mathbf{y} = (\frac{x_1}{1 - x_n}, \ldots, \frac{x_{n-1}}{1 - x_n}, 0).$$

Identifying $\mathbf{R}^{n-1}$ with $\{\mathbf{x} \in \mathbf{R}^n \ : \ x_n = 0\}$, we have $\mathbf{y} \in \mathbf{R}^{n-1}$. But

$$|y_1| + \cdots + |y_{n-1}| = |\frac{x_1}{1 - x_n}| + \cdots + |\frac{x_{n-1}}{1 - x_n}| = \frac{1}{1 - x_n}(|x_1| + \cdots + |x_{n-1}|) \leqslant 1,$$

since $|x_1| + \cdots + |x_{n-1}| \leqslant 1 - |x_n| \leqslant |1 - x_n|$. Thus $\mathbf{y} \in C_{n-1}$. By our induction assumption, then $\mathbf{y}$ is a convex combination of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_{n-1}$, say

$$\mathbf{y} = \sum_{i=1}^{n-1} (\lambda_i^+ \mathbf{e}_i + (\lambda_i^-)(-\mathbf{e}_i)), \text{ where } \lambda_i^{\pm} \geqslant 0, \ \sum_{i=1}^{n-1} (\lambda_i^+ + \lambda_i^-) = 1.$$

But note that $\mathbf{x}$ is a convex combination of $\mathbf{y}$ and $\mathbf{e}_n$:

$$\mathbf{x} = (1 - x_n)\mathbf{y} + x_n \mathbf{e}_n \text{ (keep in mind that } 0 \leqslant x_n < 1 \text{)}.$$

Thus $\mathbf{x}$ is a convex combination of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_{n-1}, \mathbf{e}_n$. Indeed

$$\mathbf{x} = \sum_{i=1}^{n-1} ((1 - x_n)\lambda_i^+ \mathbf{e}_i + (1 - x_n)\lambda_i^- (-\mathbf{e}_i)) + x_n \mathbf{e}_n.$$

(Check that this really is a convex combination of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_{n-1}, \mathbf{e}_n$.)

 The case where $-1 \leqslant x_n \leqslant 0$ can be treated similarly. This proves that any point of C_n is a convex combination of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$. In other words

$$C_n \subseteq \text{conv}\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n\}.$$

Since C_n is a convex set (Exercise 5–5) containing $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$, then C_n contains every convex combination of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$, hence

$$C_n \supseteq \text{conv}\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n\}.$$

Thus $C_n = \text{conv}\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n\}$.   ∎

Remark. It is easy to see directly the "easy part", namely that

$$C_n \supseteq \text{conv}\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n\},$$

by noting that if $\mathbf{x} = \sum_{i=1}^{n}(\lambda_i^+ \mathbf{e}_i + (\lambda_i^-)(-\mathbf{e}_i))$, where $\lambda_i^\pm \geqslant 0$, $\sum_{i=1}^{n}(\lambda_i^+ + \lambda_i^-) = 1$, then

$$\mathbf{x} = (\lambda_1^+ - \lambda_1^-, \lambda_2^+ - \lambda_2^-, \ldots, \lambda_n^+ - \lambda_n^-),$$

so

$$|x_1| + \cdots + |x_n| = |\lambda_1^+ - \lambda_1^-| + \cdots + |\lambda_n^+ - \lambda_n^-| \leqslant \lambda_1^+ + \lambda_1^- + \cdots \lambda_n^+ + \lambda_n^- = 1.$$

Hence $\mathbf{x} \in C_n$ by definition of C_n.

6.4   Exercises

6–1 If $A \subset \mathbf{R}^n$, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are convex combinations of points in A, then show that $(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2$ for $0 \leqslant \lambda \leqslant 1$ is also a convex combination of points of A.

6–2 Let $K \subseteq \mathbf{R}^3$ be the convex hull of the three points $\{(1,0,0), (0,1,0), (0,0,1)\}$. Sketch the convex polytope K.

6–3 Let $K \subseteq \mathbf{R}^4$ be the convex hull of $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$, where $\mathbf{e}_1 = (1,0,0,0)$, etc.
(a) K lies in a hyperplane. What is the equation of that hyperplane?
(b) K is a 3–dimensional convex polytope sitting in $\mathbf{R}^4$. What is the 3–dimensional volume of K?

6–4 Suppose A_1, $A_2 \subseteq \mathbf{R}^n$, with $A_1 \subseteq H_1$, and $A_2 \subseteq H_2$, where H_1 and H_2 are parallel hyperplanes (i.e. H_1 and H_2 have the same unit normal). Prove that $A_1 + A_2$ is a subset of a hyperplane parallel to H_1 and H_2.

6–5 Let K be the convex set in Exercise 6–2. Describe completely $K + (-K)$.

6–6 Let I^n be the n–dimensional cube described in Exercise 5–4. You are asked in Exercise 5–4 and Exercise 5–9 to show that I^n is convex. Show in fact that I^n is a convex polytope by proving that I^n is the convex hull of the 2^n points of the form $(x_1, \ldots, x_n)$ where each coordinate x_i is either ± 1. (For example,

$$I^2 = \text{conv}\{(1,1), (-1,1), (1,-1), (-1,-1)\}.)$$

6–7 In $\mathbf{R}^3$ let $A = \{(1,0,0), (0,1,0), (-1,0,0), (0,-1,0), (0,2,2), (0,-2,2)\}$. Let P be the convex polytope $P = \text{conv}(A)$.
(a) Sketch P and calculate its volume.
(b) Let k be the smallest integer such that each point of P is a convex combination of at most k points of A. What is k?
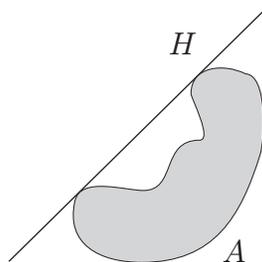
Figure 10: Set A is contained in a halfspace (Exercise 6–8).

6–8 Suppose $A \subseteq \mathbf{R}^n$ and H is a hyperplane in $\mathbf{R}^n$. Suppose A is contained in one of the halfspaces of H (Figure 10). Prove then that

$$\mathrm{conv}(H \cap A) = H \cap \mathrm{conv}(A).$$

[Hint: Suppose H has equation $\langle \mathbf{x}, \mathbf{u} \rangle = p$ and A is contained in $H^+ = \{\mathbf{x} \in \mathbf{R}^n : \langle \mathbf{x}, \mathbf{u} \rangle \geqslant p\}$. To show first that $H \cap \mathrm{conv}(A) \subseteq \mathrm{conv}(H \cap A)$, suppose $\mathbf{x} \in H \cap \mathrm{conv}(A)$, so

$$\langle \mathbf{x}, \mathbf{u} \rangle = p \text{ and } \mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i, \ \mathbf{a}_i \in A, \ \sum \lambda_i = 1, \ \lambda_i \geqslant 0.$$

Then

$$\sum_{i=1}^{k} \lambda_i (\langle \mathbf{a}_i, \mathbf{u} \rangle - p) = 0.$$

It follows from this that for each i, $i = 1, \ldots, k$, either $\lambda_i = 0$ or $\langle \mathbf{a}_i, \mathbf{u} \rangle = p$, (WHY?), so for those i such that $\lambda_i \neq 0$, $\mathbf{a}_i \in H$. Thus $\mathbf{x} \in \mathrm{conv}(H \cap A)$.

To show that $\mathrm{conv}(H \cap A) \subseteq H \cap \mathrm{conv}(A)$, suppose $\mathbf{x} \in \mathrm{conv}(H \cap A)$. Then $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{b}_i$, $\mathbf{b}_i \in H \cap A$, $\sum \lambda_i = 1$, $\lambda_i \geqslant 0$. Thus $\mathbf{x} \in \mathrm{conv}(A)$, and also $\mathbf{x} \in H$ (WHY?). That is, $\mathbf{x} \in H \cap \mathrm{conv}(A)$.]

Remark. Exercise 6–8 has a plausible "physical" interpretation. It shows that if a distribution of point masses lies in a halfspace, then the center of mass is on the boundary if and only if all the masses lie on the boundary.

6–9 Show that

$$\mathrm{conv}(A + B) = \mathrm{conv}(A) + \mathrm{conv}(B).$$

[Hint: If $\mathbf{x} = \sum_{i=1}^{k} \lambda_i (\mathbf{a}_i + \mathbf{b}_i)$, then $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i + \sum_{i=1}^{k} \lambda_i \mathbf{b}_i$. The latter two sums are convex combinations of elements of A and B, respectively. On the other hand, if $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i + \sum_{j=1}^{\ell} \mu_j \mathbf{b}_j$, with $\sum_{i=1}^{k} \lambda_i = 1$ and $\sum_{j=1}^{\ell} \mu_j = 1$, show that

$$\mathbf{x} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_i \mu_j (\mathbf{a}_i + \mathbf{b}_j). \tag{1}$$

Note that $\lambda_i \mu_j \geqslant 0$ and $\sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_i \mu_j = 1$, so the sum in (1) is a convex combination of elements from $A + B$.]

6–10 Why does Exercise 6–9 imply that the Minkowski sum of two convex polytopes is again a convex polytope.

6–11 With the n–cube defined as in Exercise 5–4, consider the six vertices of I^4 given by:

$$\mathbf{a}_1 = (-1, 1, 1, -1),\ \mathbf{a}_2 = (1, -1, 1, -1),\ \mathbf{a}_3 = (1, 1, -1, -1),\ \text{and}$$

$$\mathbf{b}_1 = (1, -1, -1, 1)\ \mathbf{b}_2 = (-1, 1, -1, 1),\ \mathbf{b}_3 = (-1, -1, 1, 1).$$

(a) Show that these six points are the vertices of a 3–dimensional regular octahedron C of edge length $2\sqrt{2}$.

[Hint: Check that the three line segments $\overline{\mathbf{a}_i\mathbf{b}_i}$, $i = 1, 2, 3$, are mutually orthogonal and intersect at their midpoints.]

(b) Note that the vertices of the octahedron C belong to the hyperplane H with equation $x_1 + x_2 + x_3 + x_4 = 0$. Explain why

$$C \subseteq H \cap I^4.$$

(c) Show also that $C \supseteq H \cap I^4$ so we have in conjunction with part (b),

$$C = H \cap I^4.$$

[Hint: If $\mathbf{x} = (x_1, x_2, x_3, x_4) \in H \cap I^4$, then $\sum_{i=1}^4 x_i = 0$ and $-1 \leqslant x_i \leqslant 1$, $i = 1, 2, 3, 4$. Note that $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ comprise an orthogonal basis for the 3–dimensional subspace H. We can assume, without loss of generality, that $\mathbf{x}$ belongs to the "positive octant" of H relative to this basis – that is, $\mathbf{x} \cdot \mathbf{a}_i \geqslant 0$, $i = 1, 2, 3$. This results in

$$\begin{aligned} -x_1 + x_2 + x_3 - x_4 &= 0, \\ x_1 - x_2 + x_3 - x_4 &= 0, \\ x_1 + x_2 - x_3 - x_4 &= 0. \end{aligned} \tag{2}$$

Define:

$$\begin{aligned} \lambda_1 &= \frac{x_2 + x_3}{2}, \\ \lambda_2 &= \frac{x_1 + x_3}{2}, \\ \lambda_3 &= \frac{x_1 + x_2}{2}. \end{aligned}$$

Show that $0 \leqslant \lambda_1 \leqslant 1$, $i = 1, 2, 3$, that $\lambda_1 + \lambda_2 + \lambda_3 \leqslant 1$, and

$$\mathbf{x} = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \lambda_3\mathbf{a}_3.$$

It follows that $\mathbf{x}$ belongs to the "positive octant" of C. A similar argument can be used for the other "octants." To show $0 \leqslant \lambda_1 \leqslant 1$, note that from

$$-x_1 + x_2 + x_3 - x_4 \geqslant 0 \text{ and } x_1 + x_2 + x_3 + x_4 = 0$$

we get

$$x_2 + x_3 = -(x_1 + x_4) \geqslant -(x_2 + x_3),$$

so $x_2 + x_3 \geqslant 0$. From this follows $\lambda_1 \geqslant 0$. Also $\lambda_1 \leqslant 1$ since $x_2 \leqslant 1$ and $x_3 \leqslant 1$. To obtain $\lambda_1 + \lambda_2 + \lambda_3 \leqslant 1$, note that

$$\lambda_1 + \lambda_2 + \lambda_3 = x_1 + x_2 + x_3 = -x_4.]$$

6–12 (a) Show that the hyperplane H in the preceding exercise is orthogonal to a principal diagonal of I^4, so the preceding exercise shows that a cross-section through the center of I^4 orthogonal to a main diagonal of the cube is a regular octahedron.

(b) What is the analogous result for I^3?

7   Open and Closed Sets

Let $S \subseteq \mathbf{R}^n$. Then S is said to be an *open set* if and only if for each $\mathbf{x} \in S$ there exists some ball centered at $\mathbf{x}$ completely contained in S.

A set $S \subseteq \mathbf{R}^n$ is said to be a *closed set* if and only if the complement of S is open. By the *complement* of S we mean

$$S^{\mathsf{c}} \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbf{R}^n : \mathbf{x} \notin S\}.$$

Example 7–A. Any open ball is an open set.
   Proof. Let B be the open ball of radius $r > 0$ centered at $\mathbf{x}_0$, so

$$B = \{\mathbf{x} \in \mathbf{R}^n : d(\mathbf{x}, \mathbf{x}_0) < r\}.$$

Let $\mathbf{x}_1 \in B$, so $d(\mathbf{x}_1, \mathbf{x}_0) = \alpha < r$. Then the ball

$$B_1 = \{\mathbf{x} \in \mathbf{R}^n : d(\mathbf{x}, \mathbf{x}_1) \leqslant \frac{1}{2}(r - \alpha)\}$$

is contained in B. For if $\mathbf{x} \in B_1$, then

$$d(\mathbf{x}, \mathbf{x}_0) \leqslant d(\mathbf{x}, \mathbf{x}_1) + d(\mathbf{x}_1, \mathbf{x}_0) \leqslant \frac{1}{2}(r - \alpha) + \alpha < r,$$

so $\mathbf{x} \in B$. ∎

Example 7–B. Any closed ball is a closed set (proof?).

Example 7–C. Any flat in $\mathbf{R}^n$ is a closed set (proof?).

Example 7–D. The empty set $\emptyset$ is both open and closed. So is $\mathbf{R}^n$.

A *boundary point* of S is a point $\mathbf{x}$ such that every ball centered at $\mathbf{x}$ contains both points belonging to S and also points not belonging to S. The *boundary* of S, denoted by ∂S, is the set of all boundary points of S. The *interior* of S consists of all points of S that are not boundary points of S. It is denoted by $\mathrm{int}(S)$. Thus we have

$$S = \partial S \cup \mathrm{int}(S) \quad \text{and} \quad \partial S \cap \mathrm{int}(S) = \emptyset.$$

Example 7–E. If B is the open ball $B = \{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) < r\}$, then the boundary of B is the sphere $\{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) = r\}$.

Example 7–F. The boundary of the closed ball $B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) \leqslant r\}$ is the sphere $\{\mathbf{x} \in \mathbf{R}^n \,:\, d(\mathbf{x}, \mathbf{x}_0) = r\}$. The interior of $B(\mathbf{x}_0, r)$ is the open ball in Example 7–E.

Example 7–G. The boundary of a halfspace (open or closed) is the hyperplane determining that halfspace.

Example 7–H. If H is a k–dimensional flat in $\mathbf{R}^n$, with $0 \leqslant k \leqslant n - 1$, then the boundary of H is H itself. (Try, for instance, the case of a line in $\mathbf{R}^2$.)

A useful criterion for closed sets is the following:

Theorem 6 *Let $S \subseteq \mathbf{R}^n$. Then S is closed if and only if every boundary point of S belongs to S.*

Proof. First, suppose S is closed. We should like to prove that S contains all its boundary points. If $\mathbf{x}$ is a boundary point of S, then every ball centered at $\mathbf{x}$ contains points of S. Since S^{C} is open, this means $\mathbf{x} \notin S^{\mathsf{C}}$. Hence $\mathbf{x} \in S$.

Conversely, suppose S contains all its boundary points. Thus if $\mathbf{x} \in S^{\mathsf{C}}$, then $\mathbf{x}$ is not a boundary point of S, hence there is a ball B centered at $\mathbf{x}$ that is either completely contained in S or completely contained in S^{C}. But B is not contained in S since $\mathbf{x} \in B$ and $\mathbf{x} \in S^{\mathsf{C}}$. Hence B is contained in S^{C}. Thus we have shown that for each $\mathbf{x} \in S^{\mathsf{C}}$ there is a ball centered at $\mathbf{x}$ completely contained in S^{C}; hence S^{C} is open, so S is closed. This completes the proof. ∎

A set $S \subseteq \mathbf{R}^n$ is said to be *bounded* if and only if it is contained in some ball.

Theorem 7 *Suppose $S \subseteq \mathbf{R}^n$ is closed and bounded. Suppose $f = f(\mathbf{x})$ is a continuous real valued function defined on S. Then f attains a maximum value on S. That is, there exists $\mathbf{x}_0 \in S$ such that $f(\mathbf{x}_0) \geqslant f(\mathbf{x})$ for all $\mathbf{x} \in S$.*
Similarly, f attains a minimum value on S.

Remark. It is useful to keep in mind that the preceding theorem is a generalization of the theorem in calculus asserting that a continuous real valued function defined on a closed interval $I \subseteq \mathbf{R}$ attains both a maximum and a minimum value on I.
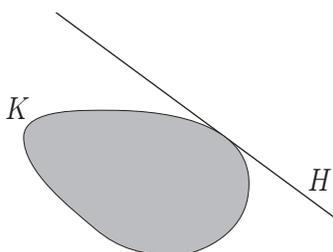
Figure 11: A supporting hyperplane H "touching" a convex set K.

8   Supporting Hyperplanes

Let K be a closed and bounded convex set in $\mathbf{R}^n$. Then a hyperplane H is called a *supporting hyperplane* of K if and only if

    (i) K is contained in one of the halfspaces of H, and

    (ii) $K \cap H \neq \emptyset$.

In other words, H is a hyperplane "touching" K, with K lying to one side of H (Figure 11). Let K be a closed and bounded convex set in $\mathbf{R}^n$, and let H be a supporting hyperplane of K. Then $H \cap K$ is called a *face* of K. Note that a face of K, being the intersection of convex sets, is again a convex set.

Example 8–A. Any face of a closed ball is a single point.

Example 8–B. The faces of a convex polytope are convex polytopes of lower dimension
(Exercise 8–1). The 0–dimensional faces are called *vertices* (singular is *vertex*), the 1–
dimensional faces *edges*, and the $(n-1)$–dimensional faces are called *facets*. The $(n-2)$–
dimensional faces of an n–dimensional polytope are often referred to as *subfacets*. For
example,

	Number of		
	vertices	edges	facets
3-dimensional cube	8	12	6
3-dimensional crosspolytope	6	12	8

Theorem 8 *Let K be a closed and bounded convex set in $\mathbf{R}^n$. Let $\mathbf{u}$ be given, with $\|\mathbf{u}\| = 1$. Then K has a supporting hyperplane H with equation $\langle \mathbf{x}, \mathbf{u} \rangle = p$, with $K \subseteq H^-$ (Figure 12). (In other words, K has a supporting hyperplane orthogonal to $\mathbf{u}$, and with $\mathbf{u}$ as "outward pointing" unit normal.)*

Proof. The function $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{u} \rangle$, $\mathbf{x} \in K$, is a continuous real valued function on the closed and bounded set K. Hence f attains a maximum value at some point $\mathbf{x}_0 \in K$. Then
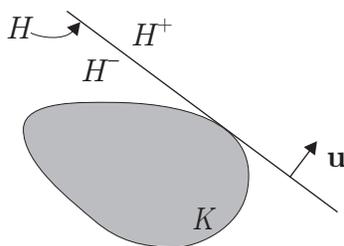
Figure 12: K has a supporting hyperplane orthogonal to $\mathbf{u}$.

$f(\mathbf{x}_0) \geqslant f(\mathbf{x})$ for all $\mathbf{x} \in K$; i.e. $\langle \mathbf{x}_0, \mathbf{u} \rangle \geqslant \langle \mathbf{x}, \mathbf{u} \rangle$, for all $\mathbf{x} \in K$. Now let H be the hyperplane with equation $\langle \mathbf{x}_0, \mathbf{u} \rangle = \langle \mathbf{x}, \mathbf{u} \rangle$. That is,

$$H = \{\mathbf{x} \in \mathbf{R}^n \, : \, \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle = 0\}.$$

Then observe that
  (i) $K \subseteq H^-$, since $\langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \rangle \leqslant 0$ for all $\mathbf{x} \in K$, and
  (ii) $K \cap H \neq \emptyset$, since $\mathbf{x}_0 \in K \cap H$.
Thus H is the required supporting hyperplane.   ∎

Example 8–C. In $\mathbf{R}^3$ let K be the crosspolytope, as defined in Exercise 5–5. Then K is a closed and bounded convex set. Let $\mathbf{u} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. The plane H with equation $\langle \mathbf{x}, \mathbf{u} \rangle = \frac{1}{\sqrt{3}}$, that is with equation

$$x + y + z = 1,$$

is a supporting plane of K with "outward normal" $\mathbf{u}$. Note that $H \cap K$ is the triangle with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, a triangular facet of K.

Remark. Let P be an n–dimensional convex polytope. Let $f_k(P)$ denote the number of k–dimensional faces of P, for $k = 0, 1, 2, \ldots, n-1$. In particular, $f_0(P)$ is the number of vertices, $f_1(P)$ is the number of edges, and $f_{n-1}(P)$ is the number of facets of P. An important relationship exists among these numbers, namely,

Euler's Formula for Polyhedra:

$$\boxed{\sum_{k=0}^{n-1}(-1)^k f_k(P) = 1 + (-1)^{n-1}.} \tag{3}$$

For example, in $\mathbf{R}^3$, for any 3–dimensional polytope we have

$$f_0(P) - f_1(P) + f_2(P) = 2.$$

We do not prove this result in these notes, but Exercises 8–12 and 8–13 provide a proof in case $n = 3$.

8.1 Exercises

8–1 Let P be a convex polytope. Prove that
(a) each face of P is a convex polytope, and
(b) P has only finitely many faces [Hint: Exercise 6–8!]

8–2 Let C_4 be the 4–dimensional crosspolytope (Exercise 5–5).
(a) How many vertices does C_4 have?
(b) Each facet of C_4 is a 3–dimensional polytope. Describe what type of 3-polytope each facet is. How many facets are there?

8–3 Is it true that every closed and bounded convex set K in $\mathbf{R}^n$ has at least one 0–dimensional face? [Hint: Consider a point of K for which $\|\mathbf{x}\|$ is maximized for $\mathbf{x} \in K$.]

8–4 Give an example of a closed and bounded convex set in $\mathbf{R}^3$ which has exactly one 1–dimensional face.

8–5 Let $A \subseteq \mathbf{R}^4$ be the set of 16 points of the form (x_1, x_2, x_3, x_4), where each x_i is either ± 1, together with the 8 points of the form $(\pm 2, 0, 0, 0)$, $(0, \pm 2, 0, 0)$, $(0, 0, \pm 2, 0)$, $(0, 0, 0, \pm 2)$. Then A is a set of 24 points in $\mathbf{R}^4$, and $P=\mathrm{conv}(A)$ is a certain convex polytope in $\mathbf{R}^4$.
(a) Explain why P is centrally symmetric with center at $\mathbf{0}$.
(b) The hyperplane H with equation $x_1 + x_2 = 2$ contains the six points $(2, 0, 0, 0)$, $(0, 2, 0, 0)$, $(1, 1, 1, 1)$, $(1, 1, -1, 1)$, $(1, 1, 1, -1)$, $(1, 1, -1, -1)$, and no other points of A. Why is $H \cap P$ the convex hull of these six points?
(c) Let $\mathbf{a}_1 = (2, 0, 0, 0)$, $\mathbf{b}_1 = (0, 2, 0, 0)$, $\mathbf{a}_2 = (1, 1, 1, 1)$, $\mathbf{b}_2 = (1, 1, -1, -1)$, $\mathbf{a}_3 = (1, 1, -1, 1)$, $\mathbf{b}_3 = (1, 1, 1, -1)$. Check that the three line segments formed by joining $\mathbf{a}_i$ to $\mathbf{b}_i$, $i = 1, 2, 3$, are of equal length, mutually orthogonal, and bisect each other. In other words they form a "3-dimensional cross". What do you conclude about the convex hull of the six points? $H \cap P$ is a facet of P of what type?
(d) Any of the 24 hyperplanes with equation of the form $\pm x_i \pm x_j = 2$, where $1 \leqslant i < j \leqslant 4$, contains exactly six points of A and intersects P in a facet congruent to that in part (c). Check this in a couple of cases. (This is the famous "24–cell", a regular polytope in $\mathbf{R}^4$ having 24 facets, where each facet is a regular octahedron. See Coxeter (1948) for a discussion of the 24–cell and other interesting regular polytopes. A picture of a projection of the 24–cell into 3–dimensional space can be found in Hilbert & Cohn-Vossen (1952, p. 152).)

8–6 If P is the 24–cell in Exercise 8–5, show that $f_0(P) = f_3(P) = 24$ (page 28) and $f_1(P) = f_2(P) = 96$. Also show that eight edges meet at each vertex, three triangles meet at each edge, and two facets meet at each 2–face.

8–7 Using Exercise 8–6, check that Euler's formula (3) holds for the 24–cell.

8–8 Show that
$$f_k(I^n) = 2^{n-k}\binom{n}{k},$$

where I^n is the n–cube (Exercise 5–4), and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

is the *binomial coefficient* "n choose k."

[Hint: Since $I^n = I^{n-1} \times I$, where $I = [-1,1]$, we see that if F_{k-1} is a $(k-1)$–dimensional face of I^{n-1}, then $F_{k-1} \times I$ is a k–dimensional face of I^n. Thus $f_k(I^n) = 2f_k(I^{n-1}) + f_{k-1}(I^{n-1})$. Use this to argue by induction. Recall that

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}.]$$

8–9 Check that Euler's formula holds for the n–cube I^n (Exercise 5–4). [Hint: By the Binomial Theorem,

$$(2-1)^n = 2^n - \binom{n}{1}2^{n-1} + \binom{n}{2}2^{n-2} - \cdots + (-1)^n.]$$

8–10 An n–*dimensional simplex* S has $n+1$ vertices. Each k–face of S is a k–dimensional simplex, and every $k+1$ vertices of S determine one of its k–faces, $k = 0, 1, \ldots, n$. Check that S satisfies Euler's formula. [Hint: By the Binomial Theorem,

$$(1-1)^{n+1} = \binom{n+1}{0} - \binom{n+1}{1} + \binom{n+1}{2} - \cdots + (-1)^{n+1}.]$$

8–11 The n–dimensional crosspolytope C_n was defined in Exercise 5–5 (see also Theorem 5, page 21).
(a) Show that
$$f_k(C_n) = 2^{k+1}\binom{n}{k+1}, \quad k = 0, 1, \ldots, n-1.$$

(b) Verify Euler's formula for C_n.

[Hint: C_n is the convex hull of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$, where $\mathbf{e}_1 = (1, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, 0, \ldots, 0, 1)$. Note that the convex hull of $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_{n-1}$ is a copy of C_{n-1} lying in the subspace $x_n = 0$, and

$$C_n = \text{conv}(-\mathbf{e}_n, \mathbf{e}_n, C_{n-1}).$$

If F is a $(k-1)$–dimensional face of C_{n-1}, then $\text{conv}(\mathbf{e}_n, F)$ and $\text{conv}(-\mathbf{e}_n, F)$ are k–faces of C_n. Thus
$$f_k(C_n) = 2f_{k-1}(C_{n-1}) + f_k(C_{n-1}).$$

this gives us a basis for induction, by showing that

$$f_k(C_n) = 2^{k+1}\binom{n}{k+1}, \ k = 0, 1, \ldots, n-1, \ \text{follows from}$$

$$f_k(C_{n-1}) = 2^{k+1}\binom{n-1}{k+1}, \ k = 0, 1, \ldots, n-2.]$$

8–12 Let $S^2 = \partial B(\mathbf{0}, 1)$ be the unit sphere in $\mathbf{R}^3$. That is

$$S^2 = \{\mathbf{u} \in \mathbf{R}^3 \ : \ \|\mathbf{u}\| = 1\}.$$

A *hemisphere* of S^2 is the intersection of S^2 with a closed halfspace of $\mathbf{R}^3$ whose bounding plane passes through $\mathbf{0}$. That is, for some $\mathbf{u}$ with $\|\mathbf{u}\| = 1$,

$$H = \{\mathbf{x} \in S^2 \ : \ \langle \mathbf{x}, \mathbf{u} \rangle \leqslant 1\}.$$

In this case we call $\mathbf{u}$ the *pole* of H.

If $H_1, H_2, \ldots, H_n$ are distinct hemispheres of S^2 such that no three of their poles lie on the same great circle, then

$$P = H_1 \cap H_2 \cap \cdots \cap H_n$$

is a *spherical convex polygon*. P is bounded by n *sides*, each being the shorter arc of a great circle. For example, a *spherical triangle* T has the form $T = H_1 \cap H_2 \cap H_3$ for distinct hemispheres H_1, H_2, H_3 not having their poles on the same great circle. T is bounded by three arcs of great circles.

(a) Suppose H_1 and H_2 are distinct hemispheres. Define real valued functions $g_i \ : \ S^2 \to \mathbf{R}$, $i = 1, 2$, by

$$g_i(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in H_i \\ 0 & \mathbf{x} \notin H_i \end{cases} \quad \text{for } \mathbf{x} \in S^2.$$

Let L be the *lune* $L = H_1 \cap H_2$, and let α be the angle of the lune, that is, the angle of intersection of the two semicircles bounding L (note that $\cos \alpha = -\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$, where $\mathbf{u}_1, \mathbf{u}_2$ are the poles of H_1, H_2, respectively).

For a region $M \subset S^2$ and $g \ : \ S^2 \to \mathbf{R}$, we let $\int\int_M g \, du$ denote the integral of g over M with respect to surface area on S^2. Show that

$$\int\int_{S^2} g_i \, du = \int\int_{S^2}(1 - g_i)\, du = 2\pi, \ i = 1, 2,$$

and also that

$$\int\int_{S^2} g_1 g_2 \, du = \int\int_{S^2}(1 - g_1)(1 - g_2)\, du = \int\int_L du = 2\alpha.$$

(b) Given a spherical triangle $T = H_1 \cap H_2 \cap H_3$ and corresponding functions g_i, $i = 1, 2, 3$, defined as in part (a), show that

$$\int\int_T du = \int\int_{S^2} g_1 g_2 g_3 \, du = \int\int_{S^2} (1 - g_1)(1 - g_2)(1 - g_3) \, du.$$

[Hint: $g_i(\mathbf{x}) + g_i(-\mathbf{x}) = 1$ if $\mathbf{x} \in S^2$ and $\mathbf{x}$ does not belong to any of the three great circles bounding H_1, H_2, H_3.]

(c) From part (b) deduce that if T is a spherical triangle with angles α, β, γ and area A, then

$$A = \alpha + \beta + \gamma - \pi.$$

Remark. By the *angle at the vertex of a spherical polygon*, we mean the measure of the angle formed by the arcs of the great circles for the sides meeting at that vertex as in the case of a lune in part (a).

[Hint: $\int\int_{S^2} g_1 \, du = 2\pi$, $\int\int_{S^2} g_1 g_2 \, du = 2\alpha$, $\int\int_{S^2} g_1 g_2 g_3 \, du = A$.]

(d) Let P be a spherical convex polygon on S^2 with n sides and vertex angles $\alpha_1, \alpha_2, \ldots, \alpha_n$. Show that the area of P is

$$A(P) = \left(\sum_{i=1}^n \alpha_i\right) - (n - 2)\pi.$$

[Hint: You may assume that P can be partitioned into $n - 2$ spherical triangles having their vertices among the vertices of P.]

8–13 Suppose P is a convex polytope in $\mathbf{R}^3$ with $\mathbf{0}$ interior to P. By the *central projection* of ∂P onto $S^2 = \partial B(\mathbf{0}, 1)$ we mean the mapping of ∂P onto S^2 that sends each $\mathbf{x} \in \partial P$ to a point $\mathbf{y} \in S^2$, where the ray through $\mathbf{x}$ emanating from $\mathbf{0}$ intersects S^2 at $\mathbf{y}$.

(a) Explain why the image of any edge of P under this mapping is an arc of a great circle having length less than π.

(b) Explain why the image of any facet of P under this mapping is a convex spherical polygon (Exercise 8–12) contained in a hemisphere of S^2.

(c) The image of ∂P under central projection induces a *network* of great circle arcs on S^2 partitioning S^2 into spherical polygons. The great circle arcs (which are images of edges of P) we shall call *edges* of the network. The points where the edges of the network meet (which are images of vertices of P) we shall call *vertices* of the network. The spherical polygons of the network (which are images of the facets of P) we shall call the *faces* of the network. If P has v vertices, e edges, and f facets, then this network clearly has v vertices, e edges, and f faces.

Let p_k be the number of spherical k–gons in the network, $k = 3, 4, \ldots$. Then $f = p_3 + p_4 + p_5 + \cdots$. Show that

$$2e = 3p_3 + 4p_4 + 5p_5 + \cdots.$$

[Hint: If we try to count the total number of edges by counting the edges of each face, each edge gets counted twice.]

Also show that

$$p_3 + 2p_4 + 3p_5 + 4p_6 + \cdots = 2(e - f).$$

(d) Let Σ be the sum of the angles of all the faces of the network on S^2 produced by the central projection of the polytope P. Show that

$$\Sigma = 2\pi v,$$

and also (using Exercise 8–12 (d)) that

$$\Sigma - \pi(p_3 + 2p_4 + 3p_5 + 4p_6 + \cdots) = 4\pi.$$

From this, and part (c) obtain Euler's formula for P:

$$v - e + f = 2.$$

8–14 Prove that every 3–dimensional convex polytope has at least one facet that is either a triangle, a quadrilateral, or a pentagon.

[Hint: Using the notation and results of the previous exercise, suppose $p_3 = p_4 = p_5 = 0$, and show then that

$$2e = 6p_6 + 7p_7 + \cdots \geqslant 6f.$$

Also, since at least three edges meet at each vertex, deduce that

$$2e \geqslant 3v.$$

From this deduce that $3e \geqslant 3(v + f)$ and obtain a contradiction to Euler's formula.]

8–15 In the hint for the previous exercise it was noted that for all 3–dimensional polytopes we must have

$$3v \leqslant 2e.$$

Show that we must also have

$$3v \geqslant e + 6.$$

[Hint: Explain why $3f \leqslant 2e$. Then, if $3v < e + 6$, we have

$$v - e + f < \frac{1}{3}e + 2 - e + \frac{2}{3}e = 2.]$$

8–16 A tetrahedron (3–dimensional simplex) has 6 edges, and a square based pyramid has 8 edges.

(a) Show that there is no 3–dimensional convex polytope with 7 edges [Hint: From the previous exercise we have

$$e + 6 \leqslant 3v \leqslant 2e.]$$

(b) Show that for each integer $n \geqslant 8$ there exists at least one 3–dimensional polytope with n edges.

9 The Support Function

9.1 Definition

Let K be a closed and bounded convex set in $\mathbf{R}^n$. The *support function* of K is a certain real valued function associated with K, defined for all $\mathbf{x} \in \mathbf{R}^n$. The definition is:

$$h(K, \mathbf{x}) \overset{\text{def}}{=} \max_{\mathbf{k} \in K} \langle \mathbf{k}, \mathbf{x} \rangle, \ \mathbf{x} \in \mathbf{R}^n.$$

First note that the maximum value of $\langle \mathbf{k}, \mathbf{x} \rangle$, for $\mathbf{k} \in K$, is attained at some point $\mathbf{k}_0 \in K$, since $\langle \mathbf{k}, \mathbf{x} \rangle$ is a continuous function of $\mathbf{k}$ and K is closed and bounded.

The support function $h(K, \mathbf{u})$ of K has an important geometrical interpretation when $\|\mathbf{u}\| = 1$. Suppose $\mathbf{k}_0 \in K$ is a point of K maximizing $\langle \mathbf{k}, \mathbf{u} \rangle$ for $\mathbf{k} \in K$, so

$$h(K, \mathbf{u}) = \max_{\mathbf{k} \in K} \langle \mathbf{k}, \mathbf{u} \rangle = \langle \mathbf{k}_0, \mathbf{u} \rangle.$$

Let H be the hyperplane with equation

$$\langle \mathbf{x}, \mathbf{u} \rangle = h(K, \mathbf{u}) = \langle \mathbf{k}_0, \mathbf{u} \rangle.$$

Then H is the supporting hyperplane of K with outward unit normal $\mathbf{u}$, since

(i) $K \subseteq H^-$ (since $\mathbf{x} \in K$ implies $\langle \mathbf{x}, \mathbf{u} \rangle \leqslant \langle \mathbf{k}_0, \mathbf{u} \rangle$ implies $\mathbf{x} \in H^-$)

(ii) $K \cap H \neq \emptyset$ (since $\mathbf{k}_0 \in K \cap H$).

But with the equation of H given in the "normal form" $\langle \mathbf{x}, \mathbf{u} \rangle = h(K, \mathbf{u}) = \langle \mathbf{k}_0, \mathbf{u} \rangle$, we know that $h(K, \mathbf{u})$ is the (signed) distance from $\mathbf{0}$ to H. Thus we have

> If $\|\mathbf{u}\| = 1$, then $h(K, \mathbf{u})$ is the (signed) distance from $\mathbf{0}$ to the supporting hyperplane of K with outward normal $\mathbf{u}$.

The picture to keep in mind is Figure 13.

Example 9–A. If $K = \{\mathbf{a}\}$, then

$$h(K, \mathbf{x}) = \max_{\mathbf{k} \in K} \langle \mathbf{k}, \mathbf{x} \rangle = \langle \mathbf{a}, \mathbf{x} \rangle = a_1 x_1 + \cdots + a_n x_n.$$

Thus $h(K, \mathbf{x})$ is a certain **linear function** in this case.

Example 9–B. If $K = B(\mathbf{0}, r)$, then $h(K, \mathbf{x}) = r\|\mathbf{x}\| = r(x_1^2 + \cdots + x_n^2)^{\frac{1}{2}}$. To see this, note that by the Cauchy-Schwarz inequality

$$\langle \mathbf{k}, \mathbf{x} \rangle \leqslant \|\mathbf{k}\| \|\mathbf{x}\| \leqslant r\|\mathbf{x}\|, \text{ if } \|\mathbf{k}\| \leqslant r.$$
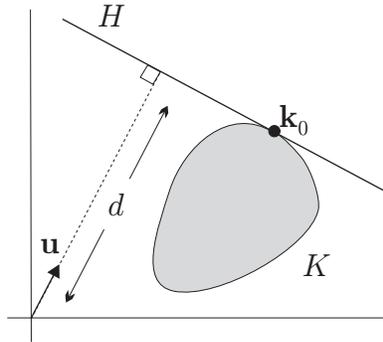
Figure 13: The supporting hyperplane H has equation $\langle \mathbf{x}, \mathbf{u} \rangle = d$, where $d = h(K, \mathbf{u})$.

But if $\mathbf{k} = \frac{r\mathbf{x}}{\|\mathbf{x}\|} \in B(\mathbf{0}, r)$ we have

$$\langle \mathbf{k}, \mathbf{x} \rangle = \langle \frac{r}{\|\mathbf{x}\|}\mathbf{x}, \mathbf{x} \rangle = \frac{r}{\|\mathbf{x}\|}\langle \mathbf{x}, \mathbf{x} \rangle = r\|\mathbf{x}\|.$$

Thus

$$h(K, \mathbf{x}) = \max_{\mathbf{k} \in K}\langle \mathbf{k}, \mathbf{x} \rangle = r\|\mathbf{x}\|.$$

Example 9–C. Suppose K is a line segment with endpoints $\mathbf{a}$ and $-\mathbf{a}$ (line segment bisected by the origin). Then

$$\mathbf{k} \in K \text{ if and only if } \mathbf{k} = (1 - \lambda)\mathbf{a} + \lambda(-\mathbf{a}) = (1 - 2\lambda)\mathbf{a}, \ 0 \leqslant \lambda \leqslant 1.$$

For fixed $\mathbf{x} \in \mathbf{R}^n$

$$\langle \mathbf{k}, \mathbf{x} \rangle = \langle (1 - 2\lambda)\mathbf{a}, \mathbf{x} \rangle = (1 - 2\lambda)\langle \mathbf{a}, \mathbf{x} \rangle.$$

But note that

$$\text{if } \langle \mathbf{a}, \mathbf{x} \rangle \geqslant 0, \text{ then } \max_{0 \leqslant \lambda \leqslant 1}(1 - 2\lambda)\langle \mathbf{a}, \mathbf{x} \rangle = \langle \mathbf{a}, \mathbf{x} \rangle, \text{ and}$$

$$\text{if } \langle \mathbf{a}, \mathbf{x} \rangle \leqslant 0, \text{ then } \max_{0 \leqslant \lambda \leqslant 1}(1 - 2\lambda)\langle \mathbf{a}, \mathbf{x} \rangle = -\langle \mathbf{a}, \mathbf{x} \rangle.$$

Thus

$$h(K, \mathbf{x}) = \max_{\mathbf{k} \in K}\langle \mathbf{k}, \mathbf{x} \rangle = |\langle \mathbf{a}, \mathbf{x} \rangle| = |a_1 x_1 + \cdots + a_n x_n|.$$

Important properties of $h(K, \mathbf{x})$:

(i)	$h(K, \mathbf{0}) = 0,$
(ii)	$h(K, \lambda\mathbf{x}) = \lambda h(K, \mathbf{x}), \text{ if } \lambda > 0,$
(iii)	$h(K, \mathbf{x} + \mathbf{y}) \leqslant h(K, \mathbf{x}) + h(K, \mathbf{y}).$

Proof of (ii). If $\lambda > 0$, then
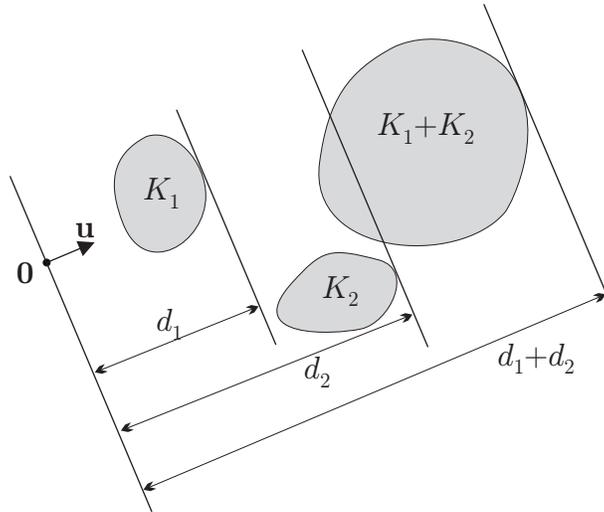
Figure 14: The support function of $K_1 + K_2$ in direction $\mathbf{u}$ is $d_1 + d_2$, Theorem 9.

$$h(K, \lambda\mathbf{x}) = \max_{\mathbf{k}\in K}\langle\mathbf{k}, \lambda\mathbf{x}\rangle = \max_{\mathbf{k}\in K}\lambda\langle\mathbf{k}, \mathbf{x}\rangle = \lambda\max_{\mathbf{k}\in K}\langle\mathbf{k}, \mathbf{x}\rangle = \lambda h(K, \mathbf{x}). \qquad \blacksquare$$

Proof of (iii).

$$\begin{aligned} h(K, \mathbf{x} + \mathbf{y}) &= \max_{\mathbf{k}\in K}\langle\mathbf{k}, \mathbf{x} + \mathbf{y}\rangle = \max_{\mathbf{k}\in K}(\langle\mathbf{k}, \mathbf{x}\rangle + \langle\mathbf{k}, \mathbf{y}\rangle) \\ &\leqslant \max_{\mathbf{k}\in K}\langle\mathbf{k}, \mathbf{x}\rangle + \max_{\mathbf{k}\in K}\langle\mathbf{k}, \mathbf{y}\rangle = h(K, \mathbf{x}) + h(K, \mathbf{y}). \end{aligned} \qquad \blacksquare$$

It is a useful and illuminating exercise to check directly that the support functions in examples A - C, above, and in the applications to ellipsoids (Section 9.3 below) satisfy these properties.

Theorem 9 *Let K_1 and K_2 be closed and bounded convex sets in $\mathbf{R}^n$. Then*

$$\boxed{h(K_1 + K_2, \mathbf{x}) = h(K_1, \mathbf{x}) + h(K_2, \mathbf{x}).}$$

Proof.

$$\begin{aligned} h(K_1 + K_2, \mathbf{x}) &= \max_{\mathbf{k}\in K_1+K_2}\langle\mathbf{k}, \mathbf{x}\rangle \\ &= \max_{\mathbf{k}_1\in K_1, \mathbf{k}_2\in K_2}\langle\mathbf{k}_1 + \mathbf{k}_2, \mathbf{x}\rangle \\ &= \max_{\mathbf{k}_1\in K_1, \mathbf{k}_2\in K_2}(\langle\mathbf{k}_1, \mathbf{x}\rangle + \langle\mathbf{k}_2, \mathbf{x}\rangle) \\ &= \max_{\mathbf{k}_1\in K_1}\langle\mathbf{k}_1, \mathbf{x}\rangle + \max_{\mathbf{k}_2\in K_2}\langle\mathbf{k}_2, \mathbf{x}\rangle \\ &= h(K_1, \mathbf{x}) + h(K_2, \mathbf{x}). \end{aligned} \qquad \blacksquare$$

See Figure 14 for a geometric interpretation.
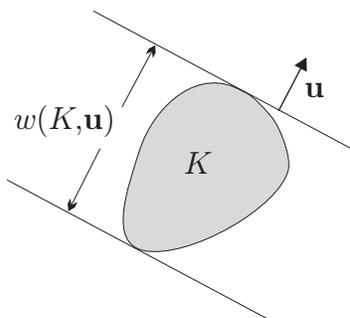
Figure 15: The width of a convex set.

9.2   The Width Function

Let $\|\mathbf{u}\| = 1$. The *width* of K in direction $\mathbf{u}$ is

$$w(K, \mathbf{u}) \stackrel{\text{def}}{=} h(K, \mathbf{u}) + h(K, -\mathbf{u}).$$

The function $w(K, \mathbf{u})$, $\|\mathbf{u}\| = 1$, is the *width function* of K. $w(K, \mathbf{u})$ is simply the distance between the supporting hyperplanes of K orthogonal to the direction $\mathbf{u}$ (Figure 15).

Theorem 10 *If $\|\mathbf{u}\| = 1$, then $h(K + (-K), \mathbf{u}) = w(K, \mathbf{u})$.*

In other words, the support function of $K + (-K)$, restricted to the unit sphere, is the width function of K.

Proof.
$$h(K + (-K), \mathbf{u}) = h(K, \mathbf{u}) + h(-K, \mathbf{u}) = h(K, \mathbf{u}) + h(K, -\mathbf{u}) = w(K, \mathbf{u}) \qquad \blacksquare$$

We say that K is a set of *constant width* if and only if $w(K, \mathbf{u}) = $ constant.

Theorem 11 *K has constant width if and only if $K + (-K)$ is a ball.*

Proof. It will later be shown in Exercise 10–6(d), that for closed and bounded convex sets in $\mathbf{R}^n$, we have $K_1 = K_2$ if and only if $h(K_1, \mathbf{u}) = h(K_2, \mathbf{u})$ for all $\mathbf{u}$ with $\|\mathbf{u}\| = 1$. If $B = B(\mathbf{0}, 1)$, this implies $K + (-K) = rB$ if and only if

$$h(K + (-K), \mathbf{u}) = h(rB, \mathbf{u}) = r \text{ for all } \mathbf{u} \text{ with } \|\mathbf{u}\| = 1.$$

That is, $K + (-K)$ is a ball of radius r if and only if $w(K, \mathbf{u}) = r$ for all $\mathbf{u}$ with $\|\mathbf{u}\| = 1$. $\blacksquare$

Example 9–D. Of course, a ball of radius r in $\mathbf{R}^n$ is a set of constant width $2r$. An example of a set of constant width that is not a ball follows (Figure 16). In $\mathbf{R}^2$, consider an equilateral triangle of side b and vertices $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. Let K be the intersection of the three circular disks of radius b centered at $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ respectively. It is a good exercise for the reader to verify that K is a plane convex set of constant width b. This particular set is known as the *Reuleaux Triangle*.
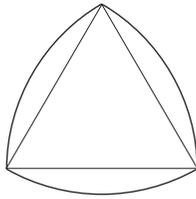
Figure 16: A Reuleaux triangle circumscribed about an equilateral triangle.

Remark. An important result which we do not prove here, is that every bounded set $A \subset \mathbf{R}^n$ is a subset of a convex body K of constant width having the same diameter as A. This is proved in Eggleston (1969, Theorem 54). The result is of great convenience in establishing results about universal covers, which we consider in § 13.1.

9.3   Ellipsoids and Superellipsoids

Example 9–E. Let E by the n–dimensional ellipsoid (Exercise 5–6),

$$E = \{\mathbf{x} \in \mathbf{R}^n \ : \ \sum_{i=1}^n \frac{x_i^2}{a_i^2} \leqslant 1\}.$$

If $\mathbf{y} \in E$, then $\sum_{i=1}^n \frac{y_i^2}{a_i^2} \leqslant 1$ and (applying the Cauchy-Schwarz inequality),

$$\langle \mathbf{y}, \mathbf{x} \rangle = \sum y_i x_i = \sum (\frac{y_i}{a_i})(a_i x_i) \leqslant (\sum \frac{y_i^2}{a_i^2})^{\frac{1}{2}} (\sum a_i^2 x_i^2)^{\frac{1}{2}} \leqslant (\sum a_i^2 x_i^2)^{\frac{1}{2}}.$$

But if we choose $y_i = \frac{a_i^2 x_i}{(\sum a_i^2 x_i^2)^{\frac{1}{2}}}$, $i = 1, \dots, n$, if $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{y} \in E$ (check this) and $\langle \mathbf{y}, \mathbf{x} \rangle = (\sum a_i^2 x_i^2)^{\frac{1}{2}}$ (check this too). Thus

$$h(E, \mathbf{x}) = \max_{\mathbf{y} \in E} \langle \mathbf{y}, \mathbf{x} \rangle = (a_1^2 x_1^2 + \cdots + a_n^2 x_n^2)^{\frac{1}{2}}.$$

This example has interesting geometrical implications. Consider the case $n = 2$, with

$$E = \{(x, y) \in \mathbf{R}^2 \ : \ \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\}.$$

With $\mathbf{x} = \mathbf{u} = (\cos \theta, \sin \theta)$ we have

$$h(E, \mathbf{u}) = (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{\frac{1}{2}}.$$

But $h(E, \mathbf{u})$ is the distance from $(0, 0)$ to the tangent line orthogonal to $(\cos \theta, \sin \theta)$, since $\|\mathbf{u}\| = 1$ (Figure 17). The boundary of E is the ellipse with equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$
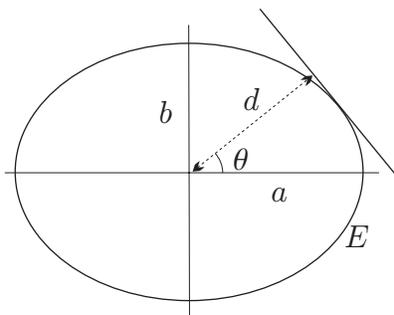
Figure 17: The ellipse in Example 9–E where $d = (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{\frac{1}{2}}$.

See Exercise 9–2 for an interesting consequence of this form of the support function of an ellipse.

Example 9–F. Now we investigate Piet Hein's Superellipsoids, see Gardner (1975, Chap. 18). Let $p > 1$, $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n > 0$ and let

$$K = \{\mathbf{x} \in \mathbf{R}^n \; : \; \frac{|x_1|^p}{a_1^p} + \cdots + \frac{|x_n|^p}{a_n^p} \leqslant 1\}.$$

Then K is a closed and bounded convex set (Exercise 9–3). K is said to be a *superellipse* when $n = 2$, and for $n > 2$ a *superellipsoid.* Using Hölder's inequality,

$$\mathbf{y} \in K \Rightarrow \sum \frac{|y_i|^p}{a_i^p} \leqslant 1 \Rightarrow$$

$$\langle \mathbf{y}, \mathbf{x} \rangle = \sum (\frac{y_i}{a_i})(a_i x_i) \leqslant (\sum \frac{|y_i|^p}{a_i^p})^{\frac{1}{p}} (\sum a_i^q |x_i|^q)^{\frac{1}{q}} \leqslant (\sum a_i^q |x_i|^q)^{\frac{1}{q}},$$

where $\frac{1}{p} + \frac{1}{q} = 1$, i.e. $q = \frac{p}{p-1}$. Let

$$\epsilon_i = \begin{cases} +1 & \text{if } x_i > 0 \\ -1 & \text{if } x_i < 0 \end{cases}$$

and define, for $\mathbf{x} \neq \mathbf{0}$,

$$y_i = \frac{\epsilon_i a_i^q |x_i|^{\frac{q}{p}}}{(\sum a_i^q |x_i|^q)^{\frac{1}{p}}}, \; i = 1, \ldots, n.$$

Then $\mathbf{y} \in K$ (check this) and $\langle \mathbf{y}, \mathbf{x} \rangle = (\sum a_i^q |x_i|^q)^{\frac{1}{q}}$ (check this too). Hence

$$h(K, \mathbf{x}) = \max_{\mathbf{y} \in K} \langle \mathbf{y}, \mathbf{x} \rangle = (\sum a_i^q |x_i|^q)^{\frac{1}{q}}.$$

Remarks on Example 9–F.
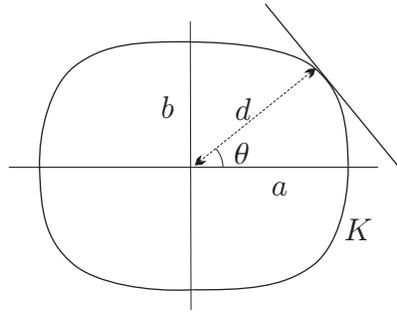
Figure 18: The superellipse in Example 9–F where $d = (a^q|\cos\theta|^q + b^q|\sin\theta|^q)^{\frac{1}{q}}$.

(i) $p = 2$ gives the case of the ellipsoid.

(ii) As $p \to 1$, note that $K \to$ crosspolytope. We would therefore expect that $h(K, \mathbf{x}) \to$ the support function of the crosspolytope. But $q \to \infty$ as $p \to 1$. Thus we need to know the value of

$$\lim_{q\to\infty} \left(\sum a_i^q |x_i|^q\right)^{\frac{1}{q}}.$$

(iii) On the other hand, as $p \to \infty$, then $K \to$ some convex set C. What is C? Since $q \to 1$ as $p \to \infty$, we expect that

$$h(C, \mathbf{x}) = \sum a_i |x_i|.$$

Make some sketches and conjectures!

(iv) Another special case to note is where $a_2 = a_3 = \cdots a_n = 0$. Then K is just the line segment joining $(-a_1, 0, \ldots, 0)$ to $(a_1, 0, \ldots, 0)$, and we obtain

$$h(K, \mathbf{x}) = a_1 |x_1|$$

in this case. But this is to be expected, in view of Example C!

(v) In case $n = 2$, with $p > 1$, and $a > b > 0$, let

$$K = \left\{(x, y) \in \mathbf{R}^2 \; : \; \frac{|x|^p}{a^p} + \frac{|y|^p}{b^p} \leqslant 1\right\}.$$

With $\mathbf{x} = \mathbf{u} = (\cos\theta, \sin\theta)$ we have then

$$h(K, \mathbf{u}) = (a^q|\cos\theta|^q + b^q|\sin\theta|^q)^{\frac{1}{q}}, \text{ where } q = \frac{p}{p-1}.$$
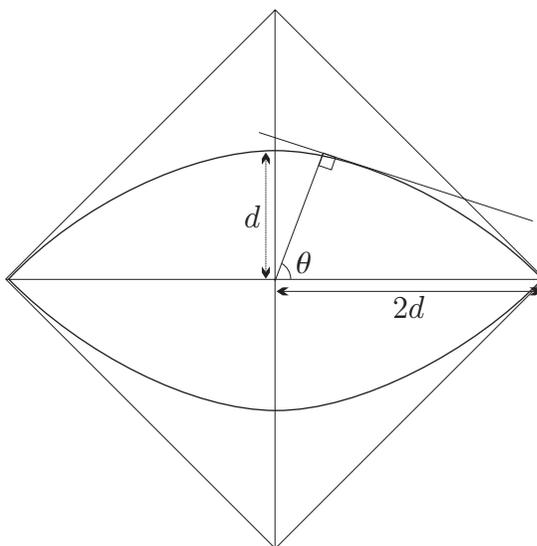
Figure 19: The "parabolic lens" for $d = \sqrt{2}/2$ (Exercise 9–1).

9.4   Exercises

9–1 Let $K \subseteq \mathbf{R}^2$ be defined by

$$K = \{(x, y) \in \mathbf{R}^2 \ : \ |y| \leqslant \frac{1}{\sqrt{2}}(1 - \frac{x^2}{2})\}.$$

(a) Verify that K is a "parabolic lens" as indicated in Figure 19.
(b) Check that for $45° \leqslant \theta \leqslant 135°$ the normal form for the tangent line to the upper parabolic arc is

$$x \cos \theta + y \sin \theta = \frac{1}{\sqrt{2} \sin \theta}.$$

(c) Show that if $\mathbf{u} = (\cos \theta, \sin \theta)$ then

$$h(K, \mathbf{u}) = \begin{cases} \frac{1}{\sqrt{2}|\sin \theta|} & 45° \leqslant \theta \leqslant 135° \text{ or } -135° \leqslant \theta \leqslant -45°, \\[2mm] \sqrt{2}|\cos \theta| & -45° \leqslant \theta \leqslant 45° \text{ or } 135° \leqslant \theta \leqslant 225°. \end{cases}$$

Describe $h(K, \mathbf{x})$ for any $\mathbf{x} = (x, y)$.
(d) Show that the circumscribed rectangles of K all have the same area.
(e) Give other examples of closed and bounded convex sets in $\mathbf{R}^2$ with the property that all the circumscribed rectangles have the same area.

9–2 (a) Using the expression for the support function of the ellipse (in $\mathbf{R}^2$) in Example 9–E, page 38, show that all the circumscribed rectangles of an ellipse have their vertices on a fixed circle (the "director circle" of the ellipse).
(b) Can you find a generalization of part (a)?
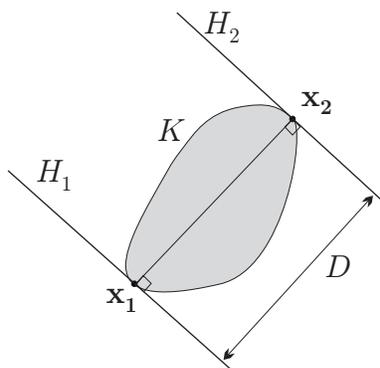
Figure 20: $D = \mathrm{diam}(K)$ (Exercise 9–8).

9–3 Let $K \subseteq \mathbf{R}^n$ be given by

$$K = \{\mathbf{x} \in \mathbf{R}^n \ : \ \frac{|x_1|^p}{a_1^p} + \cdots + \frac{|x_n|^p}{a_n^p} \leqslant 1\}, \ \text{ where } p \geqslant 1, \ a_i > 0, \ i = 1, \ldots n.$$

Prove that K is a convex set. [Hint: Use Minkowski's inequality.]

9–4 Check the details mentioned in Example 9–F, page 39.

9–5 In $\mathbf{R}^2$ let σ_1 be the line segment joining $(-a, 0)$ to $(a, 0)$, and σ_2 the line segment from $(0, -b)$ to $(0, b)$.
(a) Describe $\sigma_1 + \sigma_2$.
(b) Letting $\mathbf{x} = (x, y)$, express $h(\sigma_1, \mathbf{x})$ and $h(\sigma_2, \mathbf{x})$ as functions of x and y (Example 9–C, page 35).
(e) Give the support function of $\sigma_1 + \sigma_2$ as a function x and y.

9–6 In $\mathbf{R}^3$ let σ_i be the line segment joining $-\mathbf{e}_i$ to $\mathbf{e}_i$, $i = 1, 2, 3$, where $\mathbf{e}_1 = (1, 0, 0)$, etc.
(a) Check that $C = \sigma_1 + \sigma_2 + \sigma_3$ is the cube with vertices at the 8 points of the form $(\pm 1, \pm 1, \pm 1)$.
(b) Find $h(C, \mathbf{x})$.

9–7 This exercise, in a series of steps, will establish the continuity of the support function. This is an important property which we will need later.
Let K be a closed and bounded convex set in $\mathbf{R}^n$, and choose M such that $\|\mathbf{k}\| < M$ for all $\mathbf{k} \in K$.
(a) Show that $|h(K, \mathbf{x})| < M\|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbf{R}^n$. [Hint: By definition, $h(K, \mathbf{x}) = \max\{\langle \mathbf{k}, \mathbf{x} \rangle \ : \ \mathbf{k} \in K\}$. The Cauchy-Schwarz inequality gives $|\langle \mathbf{k}, \mathbf{x} \rangle| \leqslant \|\mathbf{k}\|\|\mathbf{x}\|$.]
(b) Show that

$$|h(K, \mathbf{x}_1) - h(K, \mathbf{x}_2)| \leqslant M\|\mathbf{x}_1 - \mathbf{x}_2\| \text{ for all } \mathbf{x}_1 \text{ and } \mathbf{x}_2 \in \mathbf{R}^n.$$

[Hint: From property (iii) on page 35 we obtain

$$
\begin{aligned}
h(K, \mathbf{x} + \mathbf{y}) - h(K, \mathbf{x}) &\leqslant h(K, \mathbf{y}), \quad \text{so} \\
h(K, \mathbf{x}_1) - h(K, \mathbf{x}_2) &\leqslant h(K, \mathbf{x}_1 - \mathbf{x}_2) \\
&\leqslant M \|\mathbf{x}_1 - \mathbf{x}_2\|.]
\end{aligned}
$$

(c) Recall that a function $f(x)$ is continuous at $\mathbf{x}_0$ if for each $\epsilon > 0$ there exists $\delta > 0$ such that
$$|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon \text{ for all } \mathbf{x} \text{ satisfying } \|\mathbf{x} - \mathbf{x}_0\| < \delta.$$

Use part (b) to show that $h(K, \mathbf{x})$, as a function of $\mathbf{x}$, is continuous at each point $\mathbf{x}_0 \in \mathbf{R}^n$. [Hint: From part (b), $|h(K, \mathbf{x}) - h(K, \mathbf{x}_0)| \leqslant M \|\mathbf{x} - \mathbf{x}_0\|$. Given $\epsilon > 0$, what choice of $\delta > 0$ will do the trick?]
Remark. (i) $h(K, \mathbf{x})$ is not only continuous, but in fact *uniformly continuous* (that is, our choice of $\delta > 0$ can be chosen independent of $\mathbf{x}_0$.
(ii) Part (b) shows that h satisfies a so-called *Lipschitz condition*.

9–8 If A is a closed and bounded subset of $\mathbf{R}^n$, then the *diameter* of A, denoted by $\text{diam}(A)$, is the maximum distance between any pair of points of A. That is

$$
\text{diam}(A) \overset{\text{def}}{=} \max_{\mathbf{x}, \mathbf{y} \in A} d(\mathbf{x}, \mathbf{y}).
$$

Suppose now that K is a closed and bounded convex set and suppose $\mathbf{x}_1$, $\mathbf{x}_2$ are a pair of points of K at a maximum distance apart, so $\text{diam}(K) = d(\mathbf{x}_1, \mathbf{x}_2)$ (Figure 20).
(a) Prove that the hyperplane H_i orthogonal to $\mathbf{x}_2 - \mathbf{x}_1$ and containing $\mathbf{x}_i$ is a supporting hyperplane of K, $i = 1, 2$.
(b) Prove that $K \cap H_i = \{\mathbf{x}_i\}$, $i = 1, 2$.
(c) Conclude that if K is a closed and bounded convex set in $\mathbf{R}^n$ and K contains more than one point, then K has at least two 0–dimensional faces.

9–9 Let K be a closed and bounded convex set in $\mathbf{R}^n$ and let $D = \text{diam}(K)$.
(a) Suppose H_1 and H_2 are parallel supporting hyperplanes of K which are distance D apart; in other words, suppose they are supporting hyperplanes orthogonal to direction $\mathbf{u}$, $\|\mathbf{u}\| = 1$, and $w(K, \mathbf{u}) = D$. Prove then that $K \cap H_i = \{\mathbf{x}_i\}$, $i = 1, 2$, with $d(\mathbf{x}_1, \mathbf{x}_2) = D$ and $\mathbf{x}_1 - \mathbf{x}_2$ orthogonal to H_1 and H_2.
(b) Prove that
$$\text{diam}(K) = \max_{\|\mathbf{u}\|=1} w(K, \mathbf{u}).$$

9–10 Let T be a triangle in $\mathbf{R}^2$.
(a) Show that $\max_{\|\mathbf{u}\|=1} w(T, \mathbf{u})$ is the length of the longest side of T. [Hint: See Exercise 9–9.]
(b) Show that $\min_{\|\mathbf{u}\|=1} w(T, \mathbf{u})$ is the length of the shortest altitude of T.

9–11 Suppose K is a set of constant width in $\mathbf{R}^n$. Prove that for each pair H_1, H_2 of parallel supporting hyperplanes of K
(a) $K \cap H_i = \{\mathbf{x}_i\}$, $i = 1, 2$, and
(b) the line segment joining $\mathbf{x}_1$ to $\mathbf{x}_2$ is orthogonal to H_1 and H_2.

9–12 Suppose K_1 and K_2 are sets of constant width in $\mathbf{R}^n$. Prove that $K_1 + K_2$ is a set of constant width.

9–13 Suppose A is a closed and bounded subset of $\mathbf{R}^n$. Prove

$$\mathrm{diam}(\mathrm{conv}(A)) = \mathrm{diam}(A).$$

9–14 Show that a fishing pole 1 mile long can be placed inside an n–dimensional cubical box with edges 1 inch long (without bending the pole), if n is sufficiently large.

9–15 Show that an elephant can be packed inside a cubical n–dimensional box with edges 1 inch long (without killing the elephant) if n is sufficiently large.

[Hint: The $4k$–dimensional cube I^{4k} has among its vertices the points $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$, where $\mathbf{x}_1$ has its first k coordinates equal to -1 and the rest 1, $\mathbf{x}_2$ has the next k coordinates -1 and all others 1, and so forth.]

10   Distance from a Point to a Set

Let K be a closed and bounded convex set in $\mathbf{R}^n$. For each $\mathbf{x}_0 \in \mathbf{R}^n$ we define "the distance from $\mathbf{x}_0$ to K" by

$$d(\mathbf{x}_0, K) \overset{\mathrm{def}}{=} \min_{\mathbf{x} \in K} d(\mathbf{x}_0, \mathbf{x}).$$

Using the fact that K is closed and bounded, and $d(\mathbf{x}_0, \mathbf{x})$ is continuous as a function of $\mathbf{x}$, it can be shown that there indeed is a point $\mathbf{x}_1 \in K$ where the minimum is achieved, so

$$d(\mathbf{x}_0, \mathbf{x}_1) = d(\mathbf{x}_0, K),$$

and in fact (using the convexity of K), $\mathbf{x}_1$ is <u>unique</u>. This point $\mathbf{x}_1$ is called *the nearest point of K to $\mathbf{x}_0$*.

Theorem 12 (Easier Supporting Theorem) *Let K be a closed and bounded convex set in $\mathbf{R}^n$ and $\mathbf{x}_0 \notin K$. Let $\mathbf{x}_1$ be the nearest point of K to $\mathbf{x}_0$. Then the hyperplane H through $\mathbf{x}_1$ orthogonal to $\mathbf{x}_0 - \mathbf{x}_1$ is a supporting hyperplane of K (in Figure 21, H has equation $\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle = 0$).*

Proof. The idea is that if there were a point $\mathbf{x}$ of K on the same side of H as $\mathbf{x}_0$, then some point of the segment $\overline{\mathbf{x}_1 \mathbf{x}}$ (which is contained in K) would be closer to $\mathbf{x}_0$ than $\mathbf{x}_1$ is.
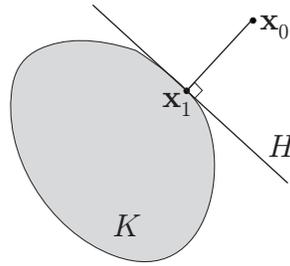
Figure 21: The Easier Supporting Theorem, Theorem 12.

In detail, since we know $\mathbf{x}_1 \in H$ it only remains to prove $K \subset H^-$ (when we write the equation of H in the form $\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle = 0$). If $\mathbf{x} \in K$ and $\mathbf{x} \notin H^-$, then we have

$$\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle = \alpha > 0.$$

The distance squared from a point $(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}$, on the segment $\overline{\mathbf{x}_1\mathbf{x}}$, to $\mathbf{x}_0$ is

$$
\begin{aligned}
\|(1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x} - \mathbf{x}_0\|^2 &= \|\mathbf{x}_1 - \mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_1)\|^2 \\
&= \|\mathbf{x}_1 - \mathbf{x}_0\|^2 - 2\lambda\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle + \lambda^2\|\mathbf{x} - \mathbf{x}_1\|^2 \\
&= \|\mathbf{x}_1 - \mathbf{x}_0\|^2 - 2\lambda\alpha + \lambda^2\|\mathbf{x} - \mathbf{x}_1\|^2.
\end{aligned}
$$

But if $0 < \lambda < \frac{2\alpha}{\|\mathbf{x}_1 - \mathbf{x}_0\|^2}$, then the r.h.s. is less than $\|\mathbf{x}_1 - \mathbf{x}_0\|^2$, giving a point of K closer than $\mathbf{x}_1$ to $\mathbf{x}_0$. This contradiction shows that $\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle \leqslant 0$, so $K \subset H^-$. ∎

Theorem 13 (Harder Supporting Theorem) *Let K be a closed and bounded convex set in $\mathbf{R}^n$ and $\mathbf{b}$ a boundary point of K. Then there exists at least one supporting hyperplane of K containing $\mathbf{b}$.*

Proof. Let S be the sphere of radius 1 centered at $\mathbf{b}$, that is

$$S = \{\mathbf{x} \in \mathbf{R}^n \ : \ d(\mathbf{x}, \mathbf{b}) = 1\}.$$

Let $\mathbf{s}_0 \in S$ be a point of S "farthest" from K, in the sense that

$$d(\mathbf{s}_0, K) = \max_{\mathbf{s} \in S} d(\mathbf{s}, K).$$

We now show that

$$d(\mathbf{s}_0, K) = d(\mathbf{s}_0, \mathbf{b}) = 1.$$

To see this, let $\epsilon > 0$ and choose $\mathbf{x}_1 \notin K$ with $d(\mathbf{x}_1, \mathbf{b}) < \epsilon$. Let H be a hyperplane through $\mathbf{x}_1$ with $K \subset H^-$ (H exists, using the previous theorem). Let H have equation

$$\langle \mathbf{x} - \mathbf{x}_1, \mathbf{u} \rangle = 0, \ \|\mathbf{u}\| = 1.$$
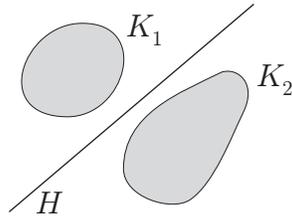
Figure 22: H "separates" K_1 and K_2.

Then $\mathbf{x} \in K$ implies:

$$
\begin{aligned}
d(\mathbf{b}+\mathbf{u},\mathbf{x}) &= \|\mathbf{b}+\mathbf{u}-\mathbf{x}\| \\
&\geqslant \langle \mathbf{b}+\mathbf{u}-\mathbf{x},\mathbf{u}\rangle, \text{ by the Cauchy-Schwarz Inequality} \\
&= \langle \mathbf{b},\mathbf{u}\rangle + \langle \mathbf{u},\mathbf{u}\rangle - \langle \mathbf{x},\mathbf{u}\rangle \\
&\geqslant \langle \mathbf{b},\mathbf{u}\rangle + \langle \mathbf{u},\mathbf{u}\rangle - \langle \mathbf{x}_1,\mathbf{u}\rangle, \text{ since } K \subset H^- \\
&= \langle \mathbf{b}-\mathbf{x}_1,\mathbf{u}\rangle + 1 \\
&\geqslant 1-\epsilon, \text{ since } |\langle \mathbf{b}-\mathbf{x}_1,\mathbf{u}\rangle| \leqslant \|\mathbf{b}-\mathbf{x}_1\| < \epsilon.
\end{aligned}
$$

Thus $d(\mathbf{b}+\mathbf{u},K) \geqslant 1-\epsilon$. Since $\mathbf{b}+\mathbf{u} \in S$, this means that $\max_{\mathbf{s}\in S} d(\mathbf{s},K) \geqslant 1-\epsilon$. Hence

$$d(\mathbf{s}_0, K) \geqslant 1-\epsilon.$$

But since this is true for any $\epsilon > 0$, we have $d(\mathbf{s}_0, K) \geqslant 1$. But note that also $d(\mathbf{s}_0, K) \leqslant 1$, since $d(\mathbf{s}_0, \mathbf{b}) = 1$. Thus

$$d(\mathbf{s}_0, K) = d(\mathbf{s}_0, \mathbf{b}) = 1,$$

as we wanted to show. Thus $\mathbf{b}$ is the nearest point of K to $\mathbf{s}_0$; hence the hyperplane through $\mathbf{b}$ orthogonal to $\mathbf{s}_0 - \mathbf{b}$ is a supporting hyperplane through $\mathbf{b}$. ∎

Theorem 14 (Separation Theorem) *Let K_1 and K_2 be a closed and bounded convex sets in $\mathbf{R}^n$ with $K_1 \cap K_2 = \emptyset$. Then there exists a hyperplane H with $K_1 \subset H^+$ and $K_2 \subset H^-$. H "separates" K_1 and K_2 (Figure 22).*

Proof. Let $\mathbf{x}_1 \in K_1$, $\mathbf{x}_2 \in K_2$ with

$$d(\mathbf{x}_1, \mathbf{x}_2) = \min_{\mathbf{k}_i \in K_i, \, i=1,2} d(\mathbf{k}_1, \mathbf{k}_2).$$

Then note that any hyperplane orthogonal to the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$, and intersecting this segment, is a separating hyperplane. ∎

Remark. It is worth noting that the "Easier Supporting Theorem" is essentially a special case of the above "Separation Theorem".

Theorem 15 *If K is a closed and bounded convex set in $\mathbf{R}^n$, then K is equal to the intersection of all its closed supporting halfspaces.*

Proof. Let M be the intersection of the supporting halfspaces of K.

If $\mathbf{x} \in K$, then $\mathbf{x}$ belongs to each supporting halfspace of K; hence $\mathbf{x} \in M$. Thus $K \subseteq M$.

If $\mathbf{x} \notin K$, then there exists a supporting halfspace of K not containing $\mathbf{x}$ (by the Easier Supporting Theorem), hence $x \notin M$. Thus $M \subseteq K$.

Thus $K = M$. ∎

Remark. This theorem is useful, since it enables one to show that a point $\mathbf{x}$ belongs to K by showing that $\mathbf{x}$ belongs to every supporting halfspace of K.

10.1 Exercises

10–1 For a convex set $K \subset \mathbf{R}^n$, we define the *polar dual* of K, denoted by K^*, by

$$K^* = \{\mathbf{y} \in \mathbf{R}^n \; : \; \langle \mathbf{x}, \mathbf{y} \rangle \leqslant 1 \text{ for all } \mathbf{x} \in K\}.$$

The following parts provide a proof that if K is a bounded convex set in $\mathbf{R}^n$ having non-empty interior, with $\mathbf{0}$ in the interior of K, then K^* is again a bounded convex set with $\mathbf{0}$ in its interior.
(a) Let $B = B(\mathbf{0}, r)$. Show that $B^* = B(\mathbf{0}, \frac{1}{r})$.
(b) Prove that if K is a convex set in $\mathbf{R}^n$, then so is K^*. [Hint: If $\langle \mathbf{x}, \mathbf{y}_1 \rangle \leqslant 1$ and $\langle \mathbf{x}, \mathbf{y}_2 \rangle \leqslant 1$, show that

$$\langle \mathbf{x}, (1 - \lambda)\mathbf{y}_1 + \lambda \mathbf{y}_2 \rangle \leqslant 1, \;\; \text{for } 0 \leqslant \lambda \leqslant 1.]$$

(c) Prove that if K_1 and K_2 are convex sets in $\mathbf{R}^n$ with $\mathbf{0} \in K_1 \subseteq K_2$, then $\mathbf{0} \in K_2^* \subseteq K_1^*$. [Hint: Why is it true that if $\langle \mathbf{y}, \mathbf{x} \rangle \leqslant 1$ for all $\mathbf{x} \in K_2$, then $\langle \mathbf{y}, \mathbf{x} \rangle \leqslant 1$ for all $\mathbf{x} \in K_1$?]
(d) If K is a bounded convex set in $\mathbf{R}^n$ having $\mathbf{0}$ in the interior of K, show that K^* is a bounded convex set having $\mathbf{0}$ as an interior point. [Hint: $\mathbf{0}$ is an interior point of K if there is a ball B_1 centered at $\mathbf{0}$ with $B_1 \subset K$, and K is bounded if there is a ball B_2 centered at $\mathbf{0}$ with $K \subset B_2$. With $B_1 \subset K \subset B_2$ we must have $B_2^* \subset K^* \subset B_1^*$.]

10–2 Suppose K is a closed and bounded convex set in $\mathbf{R}^n$, and $\mathbf{0}$ is an interior point of K. Let K^{**} denote $(K^*)^*$. Show that $K^{**} = K$.

[Hint: If $\mathbf{x} \in K$, then each $\mathbf{y} \in K^*$ satisfies $\langle \mathbf{x}, \mathbf{y} \rangle \leqslant 1$. Why does this imply $\mathbf{x} \in K^{**}$, so $K \subseteq K^{**}$? Showing $K^{**} \subseteq K$ requires more work. For this, assume $\mathbf{x}_0 \notin K$ and show that $\mathbf{x}_0 \notin K^{**}$. As in Figure 21, if $\mathbf{x}_0 \notin K$ and $\mathbf{x}_1$ is the nearest point of K to $\mathbf{x}_0$, there is a supporting hyperplane H of K through $\mathbf{x}_1$ with equation $\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle = 0$. Assuming $\mathbf{0}$ is an interior point of K, show that the equation of H can be put in the form $\langle \mathbf{x}, \mathbf{a} \rangle = 1$ for some $\mathbf{a}$. Then we have $\langle \mathbf{x}, \mathbf{a} \rangle \leqslant 1$ for all $\mathbf{x} \in K$ (implying that $\mathbf{a} \in K^*$), while $\langle \mathbf{x}_0, \mathbf{a} \rangle > 1$ (implying that $\mathbf{x}_0 \notin K^{**}$).]

10–3 Find all convex sets $K \subset \mathbf{R}^n$ with $\mathbf{0} \in K$ and $K^* = K$. [Hint: Show that if $K^* = K$, then $\langle \mathbf{x}, \mathbf{x} \rangle \leqslant 1$ for all $\mathbf{x} \in K$, so $B(\mathbf{0}, 1) \subseteq K$.]

10–4 Let $p > 1$ and $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n > 0$. Let K be the superellipsoid in $\mathbf{R}^n$ given by

$$K = \{\mathbf{x} \in \mathbf{R}^n \ : \ \frac{|x_1|^p}{a_1^p} + \cdots + \frac{|x_n|^p}{a_n^p} \leqslant 1\}.$$

Prove that if q is such that $\frac{1}{p} + \frac{1}{q} = 1$, then

$$K^* = \{\mathbf{x} \in \mathbf{R}^n \ : \ a_1^q |x_1|^q + \cdots + a_n^q |x_n|^q \leqslant 1\}.$$

[Hint: We want to find those $\mathbf{y}$ such that $\langle \mathbf{x}, \mathbf{y} \rangle \leqslant 1$ for all $\mathbf{x} \in K$. If

$$\sum_{i=1}^{n} a_i^q |y_i|^q \leqslant 1 \ \text{ and } \ \sum_{i=1}^{n} \frac{|x_i|^p}{a_i^p} \leqslant 1,$$

then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} (\frac{x_i}{a_i})(a_i y_i) \leqslant (\sum_{i=1}^{n} \frac{|x_i|^p}{a_i^p})^{\frac{1}{p}} (\sum_{i=1}^{n} a_i^q |y_i|^q)^{\frac{1}{q}} \leqslant 1,$$

showing that $\langle \mathbf{x}, \mathbf{y} \rangle \leqslant 1$ for all $\mathbf{x} \in K$, if $\mathbf{y}$ satisfies $\sum_{i=1}^{n} a_i^q |y_i|^q \leqslant 1$. On the other hand, given $\mathbf{y}$ such that $\sum_{i=1}^{n} a_i^q |y_i|^q > 1$, to construct $\mathbf{x} \in K$ such that $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} a_i^q |y_i|^q > 1$, examine the construction in Example 9–F.]

10–5 In $\mathbf{R}^n$, let I^n be the n–dimensional cube in Exercise 5–4, and C_n the n–dimensional crosspolytope in Exercise 5–5. Show that $(I^n)^* = C_n$.

[Hint: To show that $C_n \subseteq (I^n)^*$, suppose $\sum_{i=1}^{n} |y_i| \leqslant 1$ and show that $\sum_{i=1}^{n} x_i y_i \leqslant 1$ for all $\mathbf{x} \in I^n$. To show that $(I^n)^* \subseteq C_n$, suppose that $\sum_{i=1}^{n} |y_i| > 1$ (if $\mathbf{y} \notin C_n$) and find some $\mathbf{x} \in I^n$ such that $\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} |y_i| > 1$ (so $\mathbf{y} \notin (I^n)^*$).]

10–6 (a) If K is a closed and bounded convex set in $\mathbf{R}^n$ and $\mathbf{x}_0 \notin K$, let $\mathbf{x}_1$ be the nearest point of K to $\mathbf{x}_0$. Then Theorem 12, page 44, tells us that the hyperplane, H, with equation

$$\langle \mathbf{x} - \mathbf{x}_1, \mathbf{x}_0 - \mathbf{x}_1 \rangle = 0$$

is a supporting hyperplane of K through $\mathbf{x}_1$, and H separates $\mathbf{x}_0$ from K. Let $\mathbf{u} = \frac{\mathbf{x}_0 - \mathbf{x}_1}{\|\mathbf{x}_0 - \mathbf{x}_1\|}$. Show that

$$\langle \mathbf{x}_0, \mathbf{u} \rangle > h(K, \mathbf{u}).$$

[Note that $h(K, \mathbf{u}) = \langle \mathbf{x}_1, \mathbf{u} \rangle$.]
(b) From part (a) deduce that if $\mathbf{x} \in \mathbf{R}^n$ is such that $\langle \mathbf{x}, \mathbf{u} \rangle \leqslant h(K, \mathbf{u})$ for all $\mathbf{u}$ with $\|\mathbf{u}\| = 1$, then $\mathbf{x} \in K$.
(c) Use part (b) to show that if K_1 and K_2 are closed and bounded convex sets in $\mathbf{R}^n$, then

$$\boxed{K_1 \subseteq K_2 \text{ if and only if } h(K_1, \mathbf{u}) \leqslant h(K_2, \mathbf{u}) \text{ for all } \mathbf{u} \text{ with } \|\mathbf{u}\| = 1.}$$

(d) From part (c) deduce that

$$\boxed{K_1 = K_2 \text{ if and only if } h(K_1, \mathbf{u}) = h(K_2, \mathbf{u}) \text{ for all } \mathbf{u} \text{ with } \|\mathbf{u}\| = 1.}$$
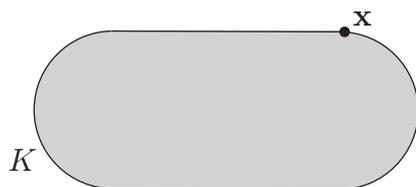
Figure 23: $\mathbf{x}$ is an extreme point of K which is not a 0–dimensional face.

11   Extreme Points

Let K be a closed and bounded convex set in $\mathbf{R}^n$. A point $\mathbf{x} \in K$ is an *extreme point* of K if and only if $\mathbf{x}$ is not "between" two points of K, i.e. there do not exist $\mathbf{x}_1$, $\mathbf{x}_2 \in K$, $\mathbf{x}_1 \neq \mathbf{x}_2$ with

$$\mathbf{x} = (1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2, \text{ with } 0 < \lambda < 1.$$

Lemma 1 *Any 0–dimensional face of K is an extreme point.*

Proof. Suppose $\mathbf{x}$ is a 0–dimensional face of K, so $\{\mathbf{x}\} = H \cap K$ for some supporting hyperplane H of K, with $K \subset H^-$. If $\mathbf{x} = (1 - \lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2$, with $\mathbf{x}_1$, $\mathbf{x}_2 \in K$, $\mathbf{x}_1 \neq \mathbf{x}_2$ and $0 < \lambda < 1$, then by Exercise 6–8 (since $\mathbf{x}_1$, $\mathbf{x}_2 \in H^-$) we must in fact have $\mathbf{x}_1$, $\mathbf{x}_2 \in H$. But that contradicts $H \cap K = \{\mathbf{x}\}$.   ∎

However, the converse of the preceding is not true (Figure 23).

Lemma 2 *Suppose F is a face of K and $\mathbf{x} \in F$. Then $\mathbf{x}$ is an extreme point of F if and only if $\mathbf{x}$ is an extreme point of K.*

Proof. See Exercise 11–1.

Examination of some examples makes plausible the following theorem of Minkowski, extended later to infinite dimensional spaces by Krein and Milman.

Theorem 16 (Krein-Milman Theorem in $\mathbf{R}^n$) *Let K be a closed and bounded convex set in $\mathbf{R}^n$, and let $\mathcal{E}$ be the set of extreme points of K. Then*

$$K = \mathrm{conv}(\mathcal{E}).$$

Proof. Since $\mathcal{E} \subseteq K$, it is clear, by Theorem 4, that $\mathrm{conv}(\mathcal{E}) \subseteq K$. What we have to prove then is that $K \subseteq \mathrm{conv}(\mathcal{E})$. In other words, we have to show that if $\mathbf{x} \in K$, then $\mathbf{x}$ is a convex combination of a finite number of extreme points $\mathbf{x}_1, \ldots \mathbf{x}_k$ of K. To do this, we proceed by induction on the dimension n. The result is obvious in dimension $n = 1$. Suppose we know it is true in all dimensions less than n, and suppose $\mathbf{x} \in K$.

If $\mathbf{x}$ is a boundary point of K, then $\mathbf{x}$ belongs to a face F of K (by the "Harder Supporting Theorem"!). By our induction hypothesis, since F is contained in a hyperplane, $\mathbf{x}$ is a convex
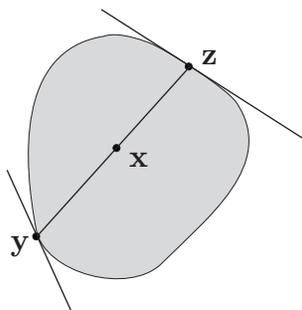
Figure 24: $\mathbf{x}$ is on a line segment with endpoints $\mathbf{y}$ and $\mathbf{z}$.

combination of extreme points of F. But each extreme point of F is an extreme point of K, by Lemma 2; hence $\mathbf{x}$ is a convex combination of extreme points of K, as required.

If $\mathbf{x}$ is not a boundary point of K, then $\mathbf{x}$ is on a line segment with endpoints $\mathbf{y}$ and $\mathbf{z}$ where $\mathbf{y}$ and $\mathbf{z}$ are boundary points of K (Why? See Figure 24.). Then $\mathbf{y}$ and $\mathbf{z}$ are each convex combinations of extreme points of K. Hence $\mathbf{x}$ is a convex combination of extreme points of K (Why?). ∎

11.1 Extreme Points of Polytopes

We will prove, in fact, that the extreme points of a polytope are precisely the vertices of the polytope. Before doing this, we prove

Theorem 17 *Let P be a convex polytope in $\mathbf{R}^n$ and let $\mathcal{V}$ be the set of vertices (i.e. 0–dimensional faces) of P. Then*
$$P = \mathrm{conv}(\mathcal{V}).$$

Proof. P is by definition the convex hull of a finite set of points of $\mathbf{R}^n$. We may assume this set is <u>minimal</u> in the sense that no element of the set is a convex combination of the others. So, suppose
$$P = \mathrm{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\},$$
where no $\mathbf{x}_i$ is a convex combination of the others.

Let $P' = \mathrm{conv}\{\mathbf{x}_2, \ldots, \mathbf{x}_k\}$. Then $\mathbf{x}_1 \notin P'$. Let $\mathbf{x}_1'$ be the nearest point of P' to $\mathbf{x}_1$ and H' the (supporting) hyperplane through $\mathbf{x}_1'$ orthogonal to $\mathbf{x}_1 - \mathbf{x}_1'$. Let H be the hyperplane through $\mathbf{x}_1$ parallel to H' (Figure 25).

Then

    (i) $P \subset H^{\pm}$, that is, P is contained in a halfspace of H, and

    (ii) $H \cap P = H \cap \mathrm{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \mathrm{conv}(H \cap \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}) = \mathrm{conv}\{\mathbf{x}_1\} = \{\mathbf{x}_1\}$.

Thus $\mathbf{x}_1 \in \mathcal{V}$. Similarly $\mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathcal{V}$. This shows that
$$\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subseteq \mathcal{V}.$$
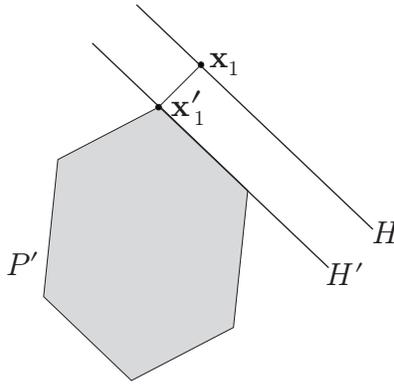
Figure 25: Illustration of the proof of Theorem 17.

On the other hand, if $\mathbf{x} \in \mathcal{V}$, then $H \cap P = \{\mathbf{x}\}$ for some supporting hyperplane of P, and so

$$\{\mathbf{x}\} = H \cap \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \mathrm{conv}(H \cap \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}),$$

so in fact $\mathbf{x} = \mathbf{x}_i$ for some i. [Note that the equation says that the convex hull of those points of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ which belong to H is just $\{\mathbf{x}\}$.] Thus

$$\mathcal{V} \subseteq \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}.$$

Thus $\mathcal{V} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ and

$$P = \mathrm{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \mathrm{conv}(\mathcal{V}). \qquad \blacksquare$$

Theorem 18 *Let P be a convex polytope in $\mathbf{R}^n$. Let $\mathcal{V}$ be the set of vertices of P and $\mathcal{E}$ the set of extreme points of P. Then*

$$\mathcal{V} = \mathcal{E}.$$

Proof. If $\mathbf{x} \in \mathcal{V}$, then since $\mathbf{x}$ is a 0–dimensional face, $\mathbf{x} \in \mathcal{E}$. Thus

$$\mathcal{V} \subseteq \mathcal{E}.$$

If $\mathbf{x} \notin \mathcal{V}$, then we have

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i, \ \lambda_i \geqslant 0, \ \sum \lambda_i = 1,$$

where $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \mathcal{V}$ and some λ_i satisfies $0 < \lambda_i < 1$. If, for example, $0 < \lambda_1 < 1$, then we can write

$$\mathbf{x} = \lambda_1 \mathbf{x}_1 + (1 - \lambda_1)(\frac{\lambda_2}{1 - \lambda_1} \mathbf{x}_2 + \cdots + \frac{\lambda_k}{1 - \lambda_1} \mathbf{x}_k),$$
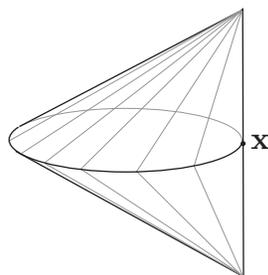
Figure 26: $\mathcal{E}$ need not be closed.

exhibiting $\mathbf{x}$ as "between" two points of P, so $\mathbf{x} \notin \mathcal{E}$. Similarly for any λ_i with $0 < \lambda_i < 1$. Thus

$$\mathcal{E} \subseteq \mathcal{V}.$$

Thus, $\mathcal{V} = \mathcal{E}$.   ∎

In the following examples $A \sim B$, read "A not B", is the set of points of A which are not in B. That is,

$$A \sim B \stackrel{\text{def}}{=} \{\mathbf{x} \ : \ \mathbf{x} \in A \text{ and } \mathbf{x} \notin B\}.$$

The *closure* of a set A, denoted by $\text{cls}(A)$, is the intersection of all closed sets that contain A. This can be shown to be the same as $A \cup \partial A$.

Example 11–A. The set of extreme points of a compact convex body need not be closed. The convex body sketched in Figure 26 is the convex hull of a circular disk and a line segment orthogonal to the disk. The point $\mathbf{x}$ is interior to the line segment and on the circumference of the disk. Since $\mathbf{x}$ is interior to the segment, $\mathbf{x} \notin \mathcal{E}$, but $\mathbf{x}$ is the limit of the points on the circumference, which do belong to $\mathcal{E}$.

Example 11–B. Klee (1958) gives an example of a 3–dimensional convex set K such that $\mathcal{E}$ and $\partial K \sim \mathcal{E}$ are both dense in ∂K!

Example 11–C. Collier (1976) proves: Let K be a convex body in $\mathbf{R}^3$. Let $\mathcal{E}$ be the set of extreme points of K. Then each component of $\text{cls}(\mathcal{E}) \sim \mathcal{E}$ is a subset of a 1–dimensional face of K (Figure 26).

11.2   Exercises

11–1 Let K be a closed and bounded convex set in $\mathbf{R}^n$. Suppose F is a face of K and $\mathbf{x} \in F$. Prove that $\mathbf{x}$ is an extreme point of F if and only if $\mathbf{x}$ is an extreme point of K. (Recall this was used in proving the Krein-Milman Theorem in $\mathbf{R}^n$.)

11–2 Let $D = \{(x, y, z) \in \mathbf{R}^3 \ : \ \sqrt{x^2 + y^2} + |z| \leqslant 1\}$.
(a) Prove that D is convex.

[Hint: Show that
$$|\lambda z_1 + (1 - \lambda)z_2| \leqslant \lambda|z_1| + (1 - \lambda)|z_2|,$$
where λ, z_1 and z_2 are real numbers and $0 \leqslant \lambda \leqslant 1$. Also use Minkowski's inequality (Exercise 2–6) to show
$$\sqrt{(\lambda x_1 + (1 - \lambda)x_2)^2 + (\lambda y_1 + (1 - \lambda)y_2)^2} \leqslant \lambda\sqrt{x_1^2 + y_1^2} + (1 - \lambda)\sqrt{x_2^2 + y_2^2},$$
where x_1, x_2, y_1 and y_2 are real numbers.]

(b) Describe the extreme points of D.

(c) Use a picture to give an "intuitive" geometric description of how D satisfies the Krein-Milman Theorem, Theorem 16.

11–3 *Radon's Theorem.*

(a) Suppose we have $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$ with $k \geqslant n + 2$. Then we know there exist $\mu_1, \ldots, \mu_k$ not all 0 with

$$\sum_{i=1}^{k} \mu_i = 0, \quad \sum_{i=1}^{k} \mu_i \mathbf{x}_i = \mathbf{0} \text{ (see Exercise 4–1 (a))}.$$

Some of the μ_i are positive and some are not. With an appropriate relabeling we may assume
$$\mu_1, \mu_2, \ldots, \mu_r > 0 \text{ and } \mu_{r+1}, \mu_{r+2}, \ldots, \mu_k \leqslant 0.$$

Then
$$\mu_1 + \cdots + \mu_r = -\mu_{r+1} - \cdots - \mu_k > 0.$$

Let
$$\begin{aligned}
\mathbf{x} &= \frac{\mu_1}{\mu_1 + \cdots + \mu_r}\mathbf{x}_1 + \cdots + \frac{\mu_r}{\mu_1 + \cdots + \mu_r}\mathbf{x}_r \\
&= -\frac{\mu_{r+1}}{\mu_1 + \cdots + \mu_r}\mathbf{x}_{r+1} - \cdots - \frac{\mu_k}{\mu_1 + \cdots + \mu_r}\mathbf{x}_k.
\end{aligned}$$

Show that
$$\mathbf{x} \in \text{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_r\} \cap \text{conv}\{\mathbf{x}_{r+1}, \ldots, \mathbf{x}_k\}.$$

(b) Prove

Theorem 19 (Radon's Theorem) *If $X \subset \mathbf{R}^n$ contains at least $n + 2$ points, then X can be partitioned into two disjoint subsets X_1 and X_2 such that*

$$\text{conv}(X_1) \cap \text{conv}(X_2) \neq \emptyset.$$

11–4 Suppose P is a convex polytope in $\mathbf{R}^n$. For each j, $j = 0, \ldots, n - 1$, prove that each point of a j–dimensional face of P is a convex combination of at most $j + 1$ vertices of P.

11–5 Let $A \subset \mathbf{R}^3$. Let $\sigma(A)$ be the set of all points $\mathbf{x} \in \mathbf{R}^3$ such that $\mathbf{x}$ belongs to a line segment with endpoints in A. Let $A_1 = \sigma(A)$ and $A_{i+1} = \sigma(A_i)$, $i = 2, \ldots$. Prove that

$$A_2 = A_3 = A_4 = \cdots.$$

11–6 Find all $\alpha > 0$ for which

$$\{\mathbf{x} \in \mathbf{R}^n : \sum_{i=1}^{n} |\mathbf{x}_i|^{\alpha} \leqslant 1\}$$

is a convex set.

11–7 Suppose K is the cube in $\mathbf{R}^n$ with vertices at the 2^n points of the form $(x_1, \ldots, x_n)$ with $x_i = 0$ or 1, $i = 1, \ldots, n$. Prove that

$$\operatorname{diam}(K) = \sqrt{n}.$$

12   Carathéodory's Theorem

If $\mathbf{x} \in \operatorname{conv}(A)$, then $\mathbf{x}$ is a convex combination of a finite number of points of A. The following theorem tells us that, in fact, in $\mathbf{R}^n$, $\mathbf{x}$ is a convex combination of no more than $n+1$ points of A, that is, $\mathbf{x}$ belongs to an at most n–dimensional simplex whose vertices are in A.

Theorem 20 (Carathéodory's Theorem) *If $A \subseteq \mathbf{R}^n$ and $\mathbf{x} \in \operatorname{conv}(A)$, then $\mathbf{x}$ is a convex combination of at most $n+1$ points of A.*

First Proof. (By induction on dimension.) It is obvious in dimension $n = 1$.

Suppose we know it is true for all dimensions less than n, and let $\mathbf{x} \in \operatorname{conv}(A)$, $A \subseteq \mathbf{R}^n$. Then $\mathbf{x} \in P = \operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ for some $\mathbf{a}_1, \ldots, \mathbf{a}_k \in A$. If $\mathbf{x}$ is a boundary point of P, let H be a supporting hyperplane of P through $\mathbf{x}$. Then

$$\mathbf{x} \in H \cap P = H \cap \operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} = \operatorname{conv}(H \cap \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}),$$

the latter equality following from Exercise 6–8. Thus $\mathbf{x}$ belongs to the convex hull of $H \cap \{\mathbf{a}_1, \ldots, \mathbf{a}_k\} \subset H$, and since H is $(n-1)$–dimensional, $\mathbf{x}$ is a convex combination of at most n points of $H \cap \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$, in particular of at most n points of A. If $\mathbf{x} \in P$, but $\mathbf{x}$ is not boundary point of P, and if $\mathbf{x} \neq \mathbf{a}_1$, then the ray $\overrightarrow{\mathbf{a}_1\mathbf{x}}$ intersects the boundary of P in a point $\mathbf{b}$. Since $\mathbf{b}$ is a convex combination of at most n points of A, then $\mathbf{x}$ is a convex combination of at most $n+1$ points of A (since $\mathbf{x}$ is on the line segment joining $\mathbf{a}_1$ to $\mathbf{b}$). ■

Second Proof. (Algebraic proof utilizing "affine dependence".) Suppose $\mathbf{x} \in \operatorname{conv}(A)$, $A \subseteq \mathbf{R}^n$. Then

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i, \text{ with } \lambda_i > 0 \text{ and } \sum_{i=1}^{k} \lambda_i = 1,$$

for some $\mathbf{a}_1, \ldots, \mathbf{a}_k \in A$. We may assume that the integer k has been chosen as small as possible in this representation. If $k \geqslant n+2$, then $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are <u>affinely dependent</u>, so there exist real numbers $\mu_1, \ldots, \mu_k$, not all zero, with

$$\sum_{i=1}^{k} \mu_i = 0 \text{ and } \sum_{i=1}^{k} \mu_i \mathbf{a_i} = \mathbf{0}$$

(Exercise 4–1(a), page 16).

Now suppose $\frac{\mu_j}{\lambda_j}$ is the largest in absolute value of the numbers $\frac{\mu_1}{\lambda_1}, \ldots, \frac{\mu_k}{\lambda_k}$. Observe that

$$\mathbf{x} = (\lambda_1 - \frac{\lambda_j}{\mu_j}\mu_1)\mathbf{a}_1 + \cdots + (\lambda_k - \frac{\lambda_j}{\mu_j}\mu_k)\mathbf{a}_k.$$

One now sees (Exercise 12–1, below) that this gives $\mathbf{x}$ as a <u>convex</u> combination of less than k elements of A, contradicting our choice of k. Hence $k < n+2$. $\blacksquare$

12.1   Exercises

12–1 In the course of the proof of Carathéodory's Theorem, given

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{a}_i, \text{ with } \lambda_i > 0 \text{ and } \sum_{i=1}^{k} \lambda_i = 1, \text{ and also}$$

$$\sum_{i=1}^{k} \mu_i \mathbf{a_i} = \mathbf{0}, \text{ with } \sum_{i=1}^{k} \mu_i = 0 \text{ and not all } \mu_i = 0, \text{ we chose}$$

an index j such that $|\frac{\mu_j}{\lambda_j}| \geqslant |\frac{\mu_i}{\lambda_i}|$ for all $i \neq j$ (why is $\mu_j \neq 0$?), and formed

$$(\lambda_1 - \frac{\lambda_j}{\mu_j}\mu_1)\mathbf{a}_1 + \cdots + (\lambda_k - \frac{\lambda_j}{\mu_j}\mu_k)\mathbf{a}_k. \tag{4}$$

(a) Show that (4) is in fact equal to $\mathbf{x}$.
(b) Show that the coefficients of $\mathbf{a}_1, \ldots, \mathbf{a}_k$ in (4) are all non-negative and sum to 1, with the coefficient of $\mathbf{a}_j$ equal to 0, so (4) indeed represents $\mathbf{x}$ as a convex combination of fewer than k elements of $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$. [Hint: Using $\lambda_i > 0$, from $|\frac{\mu_j}{\lambda_j}| \geqslant |\frac{\mu_i}{\lambda_i}|$ derive

$$\frac{\lambda_i}{\lambda_j} \geqslant \left|\frac{\mu_i}{\mu_j}\right|, \text{ and thence } \lambda_i - \frac{\lambda_j}{\mu_j}\mu_i \geqslant 0.]$$

13   Helly's Theorem and its Applications

Note that if $[a_1, b_1], \ldots, [a_n, b_n]$ are pairwise intersecting intervals in $\mathbf{R}$, then there is a point x common to them all. Indeed since $\max\{a_i : 1 \leqslant i \leqslant n\} \leqslant \min\{b_j : 1 \leqslant j \leqslant n\}$, any x satisfying

$$\max\{a_i : 1 \leqslant i \leqslant n\} \leqslant x \leqslant \min\{b_j : 1 \leqslant j \leqslant n\}$$

(in the case of closed intervals) will serve. Helly's Theorem generalizes this to families of convex sets in $\mathbf{R}^n$.

Theorem 21 (Helly's Theorem) *Given a finite number of convex sets in $\mathbf{R}^n$, suppose each $n+1$ of them have a point in common. Then they all have a point in common.*

First proof in the case of closed and bounded convex sets. (This proof is similar to Helly's original proof.) Suppose the statement of Theorem 21 is not true. Then there exists a family of closed and bounded convex sets $K_1, \ldots, K_r$ in $\mathbf{R}^n$, for some n, such that each $n+1$ intersect but not all intersect. Suppose that n has been chosen as small as possible. Then we know $n \geqslant 2$. The theorem fails for some family of r sets in $\mathbf{R}^n$, and suppose next that r has been chosen as small as possible. We have $r \geqslant n+2$ and $K_1, \ldots, K_r \subseteq \mathbf{R}^n$ such that each $n+1$ intersect and $K_1 \cap \cdots \cap K_r = \emptyset$. Moreover any $r-1$ of the K_i <u>do</u> intersect (since otherwise r would not be as small as possible).

Since $(K_1 \cap K_2 \cap \cdots \cap K_{r-1}) \cap K_r = \emptyset$, $K_1 \cap \cdots \cap K_{r-1} \neq \emptyset$, and $K_r \neq \emptyset$, there is a hyperplane H strictly separating $K_1 \cap \cdots \cap K_{r-1}$ from K_r (Theorem 14).

We claim the following: Each $r-2$ of $K_1 \cap H, \ldots, K_{r-1} \cap H$ intersect. To see this, let $K_{i_1}, \ldots, K_{i_{r-2}}$ be any $r-2$ of $K_1, \ldots, K_{r-1}$. Note that

$$(K_{i_1} \cap \cdots \cap K_{i_{r-2}}) \cap (K_1 \cap \cdots \cap K_{r-1}) = K_1 \cap \cdots \cap K_{r-1} \neq \emptyset$$

and

$$(K_{i_1} \cap \cdots \cap K_{i_{r-2}}) \cap K_r \neq \emptyset \quad (\text{``each } r-1 \text{ intersect''}).$$

Thus

$$(K_{i_1} \cap \cdots \cap K_{i_{r-2}}) \cap H \neq \emptyset \quad (H \text{ separates } K_1 \cap \cdots \cap K_{r-1} \text{ and } K_r).$$

That is,

$$(K_{i_1} \cap H) \cap \cdots \cap (K_{i_{r-2}} \cap H) \neq \emptyset$$

as we claimed.

Hence by Helly's Theorem "in H" (n was the smallest dimension where it fails!) we have

$$(K_1 \cap H) \cap \cdots \cap (K_{r-1} \cap H) \neq \emptyset,$$

i.e.

$$(K_1 \cap \cdots \cap K_{r-1}) \cap H \neq \emptyset,$$

contradicting the choice of H. This completes the proof in the case of closed and bounded convex sets.   ∎

Remark. It might help in understanding the next proof to draw a picture and follow the situation through in case $n = 2$, $r = 4$.

Second proof of Helly's Theorem. (This proof uses Radon's Theorem, Exercise 11–3, page 53, and is valid for any finite number of convex sets, not necessarily closed or bounded.

The proof is due to Radon. The proof is by induction on the number r of sets in the family in $\mathbf{R}^n$.)

The theorem is trivially true for any family of $n+1$ convex sets in $\mathbf{R}^n$. Suppose we know Helly's Theorem is true for any family of $r-1$ convex sets in $\mathbf{R}^n$, where $r \geqslant n+2$. We show then that it is true for r convex sets $K_1, \ldots, K_r$.

By hypothesis, for each $i = 1, \ldots, r$, there exists $\mathbf{x}_i$ such that

$$\mathbf{x}_i \in K_1 \cap \cdots \cap K_{i-1} \cap K_{i+1} \cap \cdots \cap K_{r-1}.$$

By Radon's Theorem, since $r \geqslant n+2$, we can partition the set of indices as

$$\{1, 2, \ldots, r\} = I \cup J$$

where $I \cap J = \emptyset$ and

$$\operatorname{conv}\{\mathbf{x}_i \,:\, i \in I\} \cap \operatorname{conv}\{\mathbf{x}_j \,:\, j \in J\} \neq \emptyset.$$

Now note that if $i \in I$, then $\mathbf{x}_i \in \bigcap_{j \in J} K_j$, and if $j \in J$, then $\mathbf{x}_j \in \bigcap_{i \in I} K_i$. Hence

$$\operatorname{conv}\{\mathbf{x}_i \,:\, i \in I\} \subseteq \bigcap_{j \in J} K_j \text{ and } \operatorname{conv}\{\mathbf{x}_j \,:\, j \in J\} \subseteq \bigcap_{i \in I} K_i.$$

Thus

$$\left(\bigcap_{j \in J} K_j\right) \cap \left(\bigcap_{i \in I} K_i\right) \neq \emptyset,$$

so $K_1 \cap \cdots \cap K_r \neq \emptyset$ as we wanted to prove. Helly's Theorem follows for any finite number of convex sets by induction. ∎

Theorem 22 *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbf{R}^2$ and suppose each three of these points can be covered by a circular disk of radius r. Then they all can be covered by a circular disk of radius r.*

Proof. Let $B = B(\mathbf{0}, 1)$ and $B_i = \mathbf{x}_i + rB$, $i = 1, \ldots, n$. Then each three of the B_i intersect (Why?). Hence there exists $\mathbf{x} \in \bigcap_{i=1}^n B_i$, by Helly's Theorem. Then $\mathbf{x} + rB$ contains all the $\mathbf{x}_i$ (Why?). ∎

Lemma 3 *If $T \subset \mathbf{R}^2$ is a triangle with all sides of length less than or equal to 1, then T can be covered by a circular disk of radius $\frac{1}{\sqrt{3}}$.*

Proof. If AB is the longest side of triangle ABC, and $d = d(A, B)$, note that C belongs to the intersection of the circular disks of radius d centered at A and B, respectively. One sees that triangle ABC is a subset of a Reuleaux Triangle of width $d \leqslant 1$ (Example 9–D), which in turn is inside a circle of radius $\frac{1}{\sqrt{3}}$. ∎

Theorem 23 (Jung's Theorem in $\mathbf{R}^2$) *If $A \subset \mathbf{R}^2$ has diameter less than or equal to 1, then A can be covered by a circular disk of radius $\frac{1}{\sqrt{3}}$.*

Proof. Each 3 points of A can be covered by a circular disk of radius $\frac{1}{\sqrt{3}}$, by Lemma 3, hence all can be covered by such a disk, by Theorem 22. ∎

Remark. The proof is incomplete since we used a theorem that was proved only for <u>finite</u> sets A. However, everything can be pushed through with a more general version of Helly's Theorem which is valid for infinite families of closed and bounded convex sets.

13.1   Universal Covers in $\mathbf{R}^2$

A convex set $S \subset \mathbf{R}^2$ is a *universal cover* if every plane set of diameter 1 can be covered by a congruent copy of S. (In other words, every set of diameter 1 is a subset of some congruent copy of S.)

Example 13–A. Jung's Theorem, Theorem 23, shows that a circular disk of radius $\frac{1}{\sqrt{3}}$ is a universal cover. Indeed, it is easy to see that it is the smallest circular universal cover.

Example 13–B. A square of side 1 is the smallest square universal cover.

Theorem 24 (Pál's Theorem) *A regular hexagon of side $\frac{1}{\sqrt{3}}$ is a universal cover.*

Proof. Let $A \subset \mathbf{R}^2$ have diameter 1. By the Remark in §9.2, page 38, A is a subset of a convex set K of constant width 1. It suffices to prove that K admits a circumscribed regular hexagon of side $\frac{1}{\sqrt{3}}$. To do this, first consider a pair of parallel supporting lines, ℓ_1 and ℓ_2, of K, in some direction. One sees that there are exactly two equilateral triangles, T_1 and T_2, each circumscribed about K, with T_1 having a side contained in ℓ_1 and T_2 a side contained in ℓ_2. (In what follows, we consider T_i as a triangular region, rather than the union of three segments.) If T_1 and T_2 happen to be the same size, then $H = T_1 \cap T_2$ is a regular hexagon circumscribed about K. (This can be deduced from the fact that T_1 and T_2 have parallel corresponding sides at distance 1 from each other.) Then H is the circumscribed regular hexagon of side $\frac{1}{\sqrt{3}}$ we want. On the other hand, if T_1 and T_2 are of different sizes, then a continuous rotation of the parallel supporting lines ℓ_1 and ℓ_2 through 180° interchanges their positions and also interchanges T_1 and T_2. But then the sizes of T_1 and T_2 have also been interchanged, so one sees by continuity that the triangles were the same size in some intermediate position, where their intersection yields the hexagon we seek. ∎

See Pál (1920) and Croft, Falconer & Guy (1991, pp. 125–127), for this theorem and related problems.

Remark. The circular disk of radius $\frac{1}{\sqrt{3}}$ is a universal cover having area $\frac{\pi}{3} \approx 1.05$, while the square of side 1 is a universal cover of area 1. The regular hexagon of side $\frac{1}{\sqrt{3}}$ is a universal cover having area $\frac{\sqrt{3}}{2} \approx 0.866$. The *Lebesgue Covering Problem* asks:
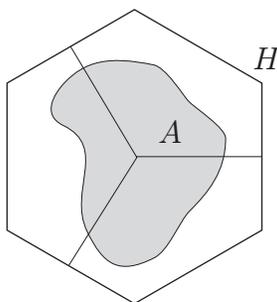
Figure 27: H partitioned into 3 sets (Theorem 25, Proof).

What is the minimum possible area of any universal cover?

This is still an unsolved problem. Exercise 13–14, below, shows that the Pál Hexagon can be truncated to yield a universal cover having area < 0.866. It is known that there exist universal covers having area at most 0.844, and that every universal cover has area at least 0.832. For a discussion of this and related questions, see Brass, Moser & Pach (2005).

13.2   Applications of Hexagons

Theorem 25 *Any plane set A of diameter 1 can be partitioned into 3 sets, each having diameter less than 1.*

Proof. Appealing to Pál's Theorem, cover A with a regular hexagon H of side $\frac{1}{\sqrt{3}}$. Note that H can be partitioned into 3 sets, each having diameter $\frac{\sqrt{3}}{2} < 1$ (Figure 27 and Exercise 12). Hence A can be partitioned into three sets of diameter less than 1. ∎

Remark. A long-standing conjecture of Borsuk, generalizing Theorem 25, was that if $A \subset \mathbf{R}^n$ has diameter 1, then A can be partitioned into $n + 1$ sets, each having diameter less than 1. This is known to be true for $n = 2$ and $n = 3$. However, as was proved by Kahn & Kalai (1993) the result is false for sufficiently large n. For more information on this, and on related unsolved problems, see Brass et al. (2005).

Theorem 26 *Suppose $D_1, \ldots, D_n$ are <u>pairwise intersecting</u> congruent circular disks in $\mathbf{R}^2$. Then there exist three points $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ such that each D_i contains at least one of $\mathbf{a}$, $\mathbf{b}$ or $\mathbf{c}$.*

Proof. Assume each D_i has radius 1, and let S be the set of centers of disks. Then $\mathrm{diam}(S) \leqslant 2$ (since each two D_i intersect). Hence S can be covered by a regular hexagon of side $\frac{2}{\sqrt{3}}$.

If $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are as indicated in Figure 28, then each point of the hexagon is within distance 1 of one of $\mathbf{a}$, $\mathbf{b}$ or $\mathbf{c}$. (Check that the circle of radius 1 centered at $\mathbf{a}$ passes through $\mathbf{b}$, $\mathbf{c}$, and two vertices of H. Similarly for $\mathbf{b}$ and $\mathbf{c}$.) Thus every disk of radius 1 with center inside the hexagon contains at least one of $\mathbf{a}$, $\mathbf{b}$ or $\mathbf{c}$. ∎
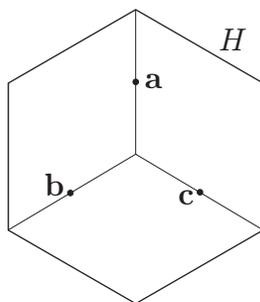
Figure 28: Each point of H is within distance 1 of one of $\mathbf{a}$, $\mathbf{b}$ or $\mathbf{c}$ (Theorem 26, Proof).

Remark. The following result, generalizing Theorem 26, and conjectured by Grünbaum (1959) (see Eckhoff (1993)), was proved by Karasev (2000):

Suppose K is a closed and bounded convex set in $\mathbf{R}^2$, and suppose $K_1, \ldots, K_n$ is a pairwise intersecting family of <u>translates</u> of K. Then there exist 3 points $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ such that each K_i contains at least one of $\mathbf{a}$, $\mathbf{b}$ or $\mathbf{c}$.

13.3   Exercises

13–1 (a) Given convex sets $K_1, \ldots K_n \subset \mathbf{R}^2$, suppose that for each three of $K_1, \ldots K_n$ there exists a circular disk of radius r intersecting all three. Prove then that there is a circular disk of radius r intersecting all of $K_1, \ldots K_n$. [Hint: Consider $K_i + rB$, $i = 1, \ldots, n$ where $B = B(\mathbf{0}, 1)$.]

(b) If each K_i is a point, what familiar result does part (a) reduce to?

13–2 Given convex sets $K_1, \ldots K_n \subset \mathbf{R}^2$, suppose that the intersection of each 3 of the K_i contains a circle of radius r. Prove then that there is a circle of radius r contained in $K_1 \cap \cdots \cap K_n$.

13–3 Let $Q = \{(x, y) \in \mathbf{R}^2 \,:\, |x| + |y| \leqslant \frac{\sqrt{2}}{2}$ and $x \leqslant \frac{1}{2}$, $y \leqslant \frac{1}{2}\}$. One observes that Q is a square circumscribed about a circle with unit diameter, with two adjacent corners truncated tangent to the circle. Prove that Q is a universal cover.

13–4 Let $Q = \{(x, y) \in \mathbf{R}^2 \,:\, x^2 + y^2 \leqslant 1, (x - 1)^2 + y^2 \leqslant 1,$ and $y \leqslant \frac{1}{2}\}$. Observe that Q is a truncated "lens". Prove that Q is a universal cover.

Remark. By the Remark on page 38, it suffices to show that Q covers every plane convex set K of constant width 1. Given such a K, from Exercise 9–11 we know that for each pair ℓ_1, ℓ_2 of parallel supporting lines of K we have $\ell_i \cap K = \mathbf{x}_i$, $i = 1, 2$, where the segment joining $\mathbf{x}_1$ to $\mathbf{x}_2$ is orthogonal to ℓ_1 and ℓ_2. Then K is contained in the "lens" L formed by intersecting the two unit disks centered at $\mathbf{x}_1$ and $\mathbf{x}_2$. The two supporting lines ℓ_3 and ℓ_4 orthogonal to ℓ_1 and ℓ_2 bound, with ℓ_1 and ℓ_2, a unit square S containing K. If the center of S coincides with the center of L, we see that

$S \cap L$ is congruent to Q. If not, use a continuity argument to show that there is some choice of ℓ_1 and ℓ_2 such that the centers do coincide.

13–5 Prove that every universal cover in $\mathbf{R}^2$ has area greater than 0.78.

13–6 A paper products manufacturing company wants to manufacture semicircular paper napkins having the property that any soup stain of diameter less than or equal to 6 inches can be covered by one of their napkins. How can they do this most economically? **Remark.** Clearly a semicircle of radius at least 6 inches is required to cover a circular disk of diameter 6. How do we know that such a semicircle will cover every set of diameter 6?

13–7 Given 300 points in $\mathbf{R}^2$, prove that there exists a point $\mathbf{p}$ such that each closed halfplane determined by each line through $\mathbf{p}$ contains at least 100 of the given points.

[Hint: Let D be a large closed disk containing all the points. For each closed halfplane H containing at least 200 points, consider $H \cap D$. This gives an infinite collection of closed and bounded convex sets. Show that each three of them intersect. Also note that if a line ℓ through a point $\mathbf{p}$ has fewer than 100 points on one side, then a slight perturbation of ℓ can be made to yield a halfplane not containing $\mathbf{p}$ and containing at least 200 points of the given set.]

13–8 Suppose P is a convex polygon in $\mathbf{R}^2$ with $\mathbf{0}$ in the interior. If $\mathbf{u} = (\cos\theta, \sin\theta)$, we abbreviate the support function of P by writing

$$h(P, \theta) = h(P, \mathbf{u}).$$

If $L(P)$ denotes the perimeter of P, prove that

$$L(P) = \int_0^{2\pi} h(P, \theta)\, d\theta.$$

[Hints: (a) In Figure 29, α_i denotes the angle from the positive x–axis to the perpendicular from the origin to the i–edge of the polygon. For r_i and θ_i as shown in the figure, and $\alpha_i \leqslant \theta \leqslant \alpha_{i+1}$, show that

$$h(P, \theta) = r_i \cos(\theta_i - \theta).$$

(b) Show that

$$\int_{\alpha_i}^{\alpha_{i+1}} h(P, \theta)\, d\theta = r_i \sin(\theta_i - \alpha_i) + r_i \sin(\alpha_{i+1} - \theta_i).$$

(c) Then

$$\int_0^{2\pi} h(P, \theta)\, d\theta = \sum_{edges} r_i \sin(\theta_i - \alpha_i) + r_i \sin(\alpha_{i+1} - \theta_i).$$

Show that the right-hand side is the perimeter of P.
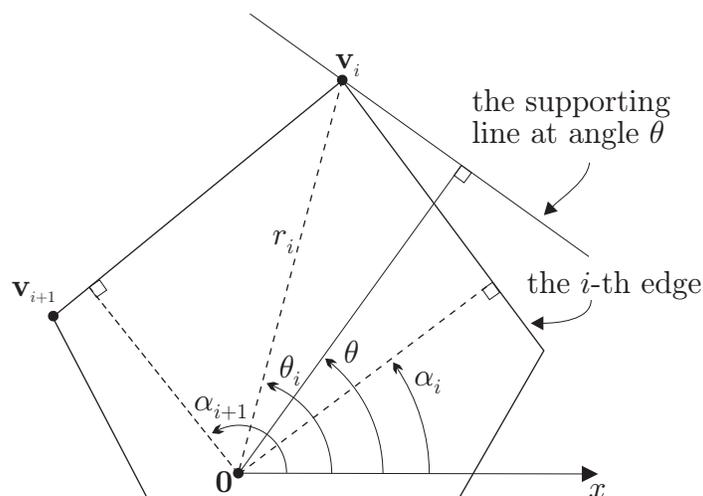
Figure 29: In Exercise 13–8 the i-th vertex of the polygon is $\mathbf{v_i}$, $\|\mathbf{v_i}\| = r_i$.

13–9 By using the result in Exercise 13–8 and the fact that

$$h(\mathbf{a} + K, \mathbf{u}) = \langle \mathbf{a}, \mathbf{u} \rangle + h(K, \mathbf{u}),$$

show that for any convex polygon P in $\mathbf{R}^2$

$$L(P) = \int_0^{2\pi} h(P, \theta)\, d\theta.$$

(In other words, the origin need not be interior to P.)

13–10 If P and Q are convex polygons in $\mathbf{R}^2$, prove that

$$L(P + Q) = L(P) + L(Q).$$

[Hint: Use Exercise 13–9.]

13–11 The integral formula for perimeter in Exercise 13–9 is actually valid for any closed and bounded plane convex set K, as can be shown by approximation arguments. In other words,

$$L(K) = \int_0^{2\pi} h(K, \theta)\, d\theta. \tag{5}$$

Assuming (5), prove (a) – (d):
(a)

$$L(K) = \frac{1}{2} \int_0^{2\pi} w(K, \theta)\, d\theta, \tag{6}$$

where $w(K, \theta)$ is the width function of K.
(b)

Theorem 27 (Barbier's Theorem) *All plane convex sets of constant width b have the same perimeter πb.*

(c)
$$L \leqslant \pi D,$$

where L and D are the perimeter and diameter of K, respectively.

(d) If K has perimeter L, then some equilateral triangle circumscribed about K has side length $\frac{\sqrt{3}}{\pi}L$.

13–12 Show that the three (irregular) pentagons in the partition of H in Figure 27 have diameter $\frac{\sqrt{3}}{2}$. [Hint: Use the fact that the diameter of a polygon is attained as the distance between some pair of its vertices.]

13–13 Show that Theorem 23 is a simple and direct consequence of Theorem 24.

13–14 (a) Let $H = ABCDEF$ be a regular hexagon of side $\frac{1}{\sqrt{3}}$ circumscribed about a convex set K of constant width 1. Let ℓ be the line closer to the vertex A which is perpendicular to the diagonal $\overline{AD}$ and tangent to the inscribed circle of H. Let H' be the "truncated hexagon" obtained by cutting off a corner of H with the line ℓ. Show that H' is a universal cover.

[Hint: Suppose H is circumscribed about a set K of constant width 1. If there are points of K inside the corner piece of H that we truncated with line ℓ, show that there are then no points of K inside the opposite corner of H obtained by similarly truncating with a line parallel to ℓ.]

(b) Show that the polygon obtained by truncating two adjacent corners of H, each in the manner of part (a), is a universal cover.

13–15 Show that if K_1 and K_2 are closed and bounded plane convex sets with $K_1 \subseteq K_2$, then $L(K_1) \leqslant L(K_2)$, and $L(K_1) = L(K_2)$ only if $K_1 = K_2$.

[Hint: Use Exercise 10–6 (c) and (d), page 48, and (5), page 62, to establish

$$L(K_1) \leqslant L(K_2) \text{ if } K_1 \subseteq K_2.$$

Next show that if $K_1 \neq K_2$, then for some θ_0 we have $h(K_1, \theta_0) - h(K_2, \theta_0) > 0$. We have shown, Exercise 9–7, page 42, that $h(K, \theta)$ is a continuous function of θ. Thus we must have $h(K_1, \theta) - h(K_2, \theta) > 0$ for θ varying over some interval centered at θ_0. Then

$$\int_0^{2\pi} h(K_1, \theta) - h(K_2, \theta)\, d\theta > 0.]$$
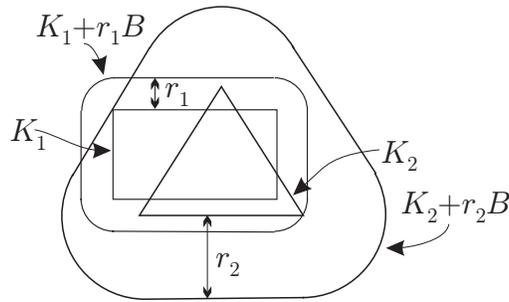
Figure 30: $K_2 \subseteq K_1 + r_1 B$ and $K_1 \subseteq K_2 + r_2 B$

14   The Distance between Convex Sets

We now define a notion of "distance" between convex sets in $\mathbf{R}^n$. In other words, we define a <u>metric</u>, or distance function, $d(K_1, K_2)$ for closed and bounded convex sets K_1, K_2 in $\mathbf{R}^n$, having the fundamental properties:

> (i) $d(K_1, K_2) \geqslant 0$, and $d(K_1, K_2) = 0$ if and only if $K_1 = K_2$,
> (ii) $d(K_1, K_2) = d(K_2, K_1)$,
> (iii) $d(K_1, K_2) + d(K_2, K_3) \geqslant d(K_1, K_3)$ (The Triangle Inequality).

Once we have done this, we will have made the set of all closed and bounded convex sets in $\mathbf{R}^n$ into a *metric space*. (See page 65 for the proof that $d(K_1, K_2)$ satisfies the three properties above.)

Let K_1, K_2 be closed and bounded convex sets in $\mathbf{R}^n$. We define *the distance from K_1 to K_2*, denoted $d(K_1, K_2)$, by

$$d(K_1, K_2) \overset{\text{def}}{=} \min\{r \geqslant 0 \ : \ K_2 \subseteq K_1 + rB, \ K_1 \subseteq K_2 + rB, \ \text{where } B = B(\mathbf{0}, 1)\}.$$

Remark. $K + rB$ is the *outer parallel set of K at distance r*. It is the union of all balls of radius r centered in K. Alternatively, one can see that $K + rB$ is the set of all points $\mathbf{x} \in \mathbf{R}^n$ whose distance from K is less than or equal to r. Thus $d(K_1, K_2)$ is the smallest r such that the outer parallel set at distance r of each set contains the other. In Figure 30,

$$r_1 = \min\{r \geqslant 0 \ : \ K_2 \subseteq K_1 + rB\} \text{ and } r_2 = \min\{r \geqslant 0 \ : \ K_1 \subseteq K_2 + rB\},$$

then

$$d(K_1, K_2) = \max\{r_1, r_2\}.$$

Example 14–A. If $K_1 = \{\mathbf{x}_1\}$ and $K_2 = \{\mathbf{x}_2\}$, then $d(K_1, K_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|$.

Example 14–B. If B_i is the ball of radius r_i centered at $\mathbf{a}_i$, $i = 1, 2$, then

$$d(B_1, B_2) = \|\mathbf{a}_1 - \mathbf{a}_2\| + |r_1 - r_2|.$$
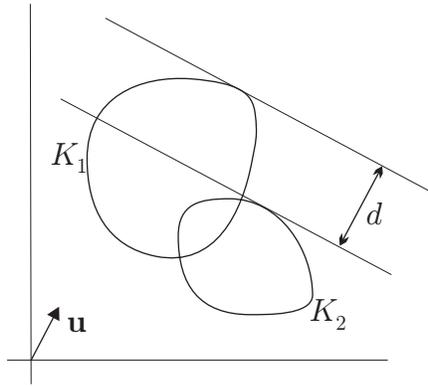
Figure 31: $d = |h(K_1, \mathbf{u}) - h(K_2, \mathbf{u})|$

The following lemma gives an interpretation of the distance function between convex sets in terms of their support functions.

Lemma 4 *If K_1, K_2 are closed and bounded convex sets in $\mathbf{R}^n$, then*

$$d(K_1, K_2) = \max_{\|\mathbf{u}\|=1} |h(K_1, \mathbf{u}) - h(K_2, \mathbf{u})|.$$

Proof. Figure 31 illustrates the following:

$$K_2 \subseteq K_1 + rB \text{ if and only if}$$

$$h(K_2, \mathbf{x}) \leqslant h(K_1 + rB, \mathbf{x}) = h(K_1, \mathbf{x}) + r\|\mathbf{x}\|, \text{ for all } \mathbf{x}, \text{ and}$$

$$K_1 \subseteq K_2 + rB \text{ if and only if } h(K_1, \mathbf{x}) \leqslant h(K_2, \mathbf{x}) + r\|\mathbf{x}\|, \text{ for all } \mathbf{x}.$$

Thus $d(K_1, K_2)$ is the smallest r such that

$$|h(K_1, \mathbf{x}) - h(K_2, \mathbf{x})| \leqslant r\|\mathbf{x}\|, \text{ for all } \mathbf{x}.$$

This is the smallest r such that $|h(K_1, \mathbf{u}) - h(K_2, \mathbf{u})| \leqslant r$ for all $\|\mathbf{u}\| = 1$. ∎

Proof that $d(K_1, K_2)$ is a metric.

(i) Clearly $d(K_1, K_2) \geqslant 0$, and $d(K_1, K_2) = 0$
    if and only if $h(K_1, \mathbf{u}) = h(K_2, \mathbf{u})$, for all $\|\mathbf{u}\| = 1$,
    if and only if $K_1 = K_2$.

(ii) Clearly $d(K_1, K_2) = d(K_2, K_1)$.

(iii) By Lemma 4

$$
\begin{aligned}
d(K_1, K_3) &= \max_{\|\mathbf{u}\|=1} |h(K_1, \mathbf{u}) - h(K_3, \mathbf{u})| \\
&= \max_{\|\mathbf{u}\|=1} |h(K_1, \mathbf{u}) - h(K_2, \mathbf{u}) + h(K_2, \mathbf{u}) - h(K_3, \mathbf{u})| \\
&\leqslant \max_{\|\mathbf{u}\|=1} |h(K_1, \mathbf{u}) - h(K_2, \mathbf{u})| + \max_{\|\mathbf{u}\|=1} |h(K_2, \mathbf{u}) - h(K_3, \mathbf{u})| \\
&= d(K_1, K_2) + d(K_2, K_3).
\end{aligned}
$$

∎

14.1   Definition of Convergence of Convex Sets

Given a sequence $K_1, K_2, K_3, \ldots$ of closed and bounded convex sets in a Euclidean space, we shall say that the sequence *converges* to a closed and bounded convex set K if and only if

$$d(K_i, K) \to 0 \text{ as } i \to \infty.$$

We shall denote this by $K_i \to K$, or $\lim_{i \to \infty} K_i = K$.

Remark. $K_i \to K$ if and only if for each $\epsilon > 0$ there exists N such that

$$K_i + \epsilon B \supseteq K \text{ and } K + \epsilon B \supseteq K_i \text{ for all } i > N.$$

Theorem 28 $K_i \to K$ *if and only if for each $\epsilon > 0$ there exists N such that $|h(K_i, \mathbf{u}) - h(K, \mathbf{u})| < \epsilon$, for all $\|\mathbf{u}\| = 1$ and for all $i > N$.*

Remark. The condition expressed in the theorem is that the sequence of functions $\{h(K_i, \mathbf{u})\}$ <u>converges uniformly</u> to $h(K, \mathbf{u})$ on the unit sphere $\{\mathbf{u} \in \mathbf{R}^n : \|\mathbf{u}\| = 1\}$.

Example 14–C. If $K_i = \{\mathbf{x}_i\}$, $i = 1, 2, \ldots$, then convergence of $\{K_i\}$ is the same as convergence of a sequence of points in the usual sense.

Example 14–D. If B_i is a ball of radius r_i centered at $\mathbf{a}_i$, $i = 1, 2, \ldots$, and if $B_i \to K$, then K is either a ball or a point.

Example 14–E. If m is a fixed integer and each K_i is a polytope with at most m vertices, and if $K_i \to K$, then K is a polytope with at most m vertices. (Note, however, that if we place no bound on the number of vertices, then the limit of a sequence of polytopes need not be a polytope!)

Remark. The proofs of the results in Examples 14–A and 14–C are very elementary exercises for the reader. We assign Example 14–B as Exercise 14–3 and Example 14–D as Exercise 14–4. However we do not prove the result of Example 14–E in these notes, but refer the reader to Grünbaum (2003, §5.3) for this and more general results about the limits of sequences of polytopes.

Theorem 29 (The Blaschke Selection Principle) *Suppose $K_1, K_2, \ldots$ is an infinite sequence of closed and bounded convex sets in $\mathbf{R}^n$, all contained in some fixed ball. Then there exists a subsequence $K_{i_1}, K_{i_2}, \ldots$ converging to a closed and bounded convex set K.*

Remark. This may be viewed as a generalization of the Bolzano-Weierstrass Theorem, which asserts that a bounded sequence of points has a convergent subsequence. A proof of the Blaschke Selection Principle is given in Schneider (1993, pp. 49–50).

Given a sequence $\{K_i\}$ of closed and bounded convex sets in $\mathbf{R}^n$, the condition that the sequence converge to K, stated as $d(K_i, K) \to 0$ as $i \to \infty$, is more precisely defined as follows:

The sequence $\{K_i\}$ converges to K if and only if for each $\epsilon > 0$ there exists N such that

$$d(K_i, K) < \epsilon \text{ for all } i \geqslant N.$$

We say that $\{K_i\}$ is a *Cauchy sequence* if it satisfies the following *Cauchy criterion*:

> For each $\epsilon > 0$ there exists N such that $d(K_i, K_j) < \epsilon$ for all $i, j \geqslant N$.

Theorem 30 *Any convergent sequence is a Cauchy sequence.*

Proof. Let $K_1, K_2, \ldots$ be a sequence such that $K_i \to K$ as $i \to \infty$. If $\epsilon > 0$ is given, choose N such that
$$d(K_i, K) < \frac{\epsilon}{2} \text{ for all } i \geqslant N.$$
Then, if $i, j \geqslant N$, we have

$$d(K_i, K_j) \leqslant d(K_i, K) + d(K, K_j) \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus the sequence satisfies the Cauchy criterion.   ∎

So every convergent sequence is a Cauchy sequence. But is every Cauchy sequence a convergent sequence?

A metric space is said to be *complete* if every Cauchy sequence is a convergent sequence. It can be shown from the Blaschke Selection Principle that the metric space of closed and bounded convex sets in $\mathbf{R}^n$ we are discussing is in fact complete. That is, if $\{K_i\}$ is a Cauchy sequence of closed and bounded convex sets in $\mathbf{R}^n$, then there exists a closed and bounded convex set K such that $K_i \to K$ as $i \to \infty$.

14.2   Exercises

14–1 If K is a closed and bounded convex set in $\mathbf{R}^n$, $B = B(\mathbf{0}, 1)$, and $r > 0$, prove that

$$K + rB = \{\mathbf{x} \in \mathbf{R}^n \ : \ d(\mathbf{x}, K) \leqslant r\}.$$

14–2 Given closed and bounded convex sets K and L in $\mathbf{R}^n$, let K_r and L_r denote the outer parallel sets of K and L, respectively, at distance $r \geqslant 0$ (Remark, page 64). Show that

$$d(K_r, L_r) = d(K, L).$$

14–3 Prove the result in Example 14–B. [Hint:$h(B(\mathbf{a}, r), \mathbf{u}) = h(B(\mathbf{0}, r) + \mathbf{a}, \mathbf{u}) = r + \langle \mathbf{a}, \mathbf{u} \rangle$.]

14–4 Prove the result in Example 14–D.

[Hint: If $B_i \to K$, then $\{B_i\}$ is a Cauchy sequence. Deduce from this that $\{\mathbf{a}_i\}$ is a Cauchy sequence in $\mathbf{R}^n$ and $\{r_i\}$ is a Cauchy sequence in $\mathbf{R}$. Since $\mathbf{R}^n$ is a complete metric space for $n = 1, 2, 3, \ldots$, there exist $\mathbf{a}_0 \in \mathbf{R}^n$ and $r_0 \in \mathbf{R}$ such that

$$\mathbf{a}_i \to \mathbf{a}_0 \text{ as } i \to \infty \text{ and } r_i \to r_0 \text{ as } i \to \infty.$$

Show then that $B_i \to B(\mathbf{a}_0, r_0)$ as $i \to \infty$.]

14–5 If $\{K_i\}$ and $\{L_i\}$ are sequences of closed and bounded convex sets in $\mathbf{R}^n$, with $K_i \to K$ as $i \to \infty$ and $L_i \to L$ as $i \to \infty$, show that

$$K_i + L_i \to K + L \text{ as } i \to \infty.$$

[Hint: Show $\max_{\|\mathbf{u}\|=1} |h(K_i + L_i, \mathbf{u}) - h(K + L, \mathbf{u})|$

$$\leqslant \max_{\|\mathbf{u}\|=1} |h(K_i, \mathbf{u}) - h(K, \mathbf{u})| + \max_{\|\mathbf{u}\|=1} |h(L_i, \mathbf{u}) - h(L, \mathbf{u})|.]$$

14–6 Suppose $\{K_i\}$ is a sequence of closed and bounded convex sets of constant width in $\mathbf{R}^n$, with $K_i \to K$ as $i \to \infty$. Prove then that K has constant width (or is a point).

[Hint: Show that $-K_i \to -K$ as $i \to \infty$. Then from the previous exercise we have $K_i + (-K_i) \to K + (-K)$ as $i \to \infty$.]

15   Approximation Theorems

Theorem 31 *Suppose K is a closed and bounded convex set in $\mathbf{R}^n$, and $\epsilon > 0$. Then there exist convex polytopes P, Q with $P \subseteq K \subseteq Q$ and $d(P, K) \leqslant \epsilon$ and $d(Q, K) \leqslant \epsilon$.*

Proof. (a) To construct Q, cover K with finitely many cubes of edge length $\frac{\epsilon}{\sqrt{n}}$. The convex hull of the set of all vertices of these cubes is a convex polytope Q with $K \subseteq Q$. Since each vertex of each cube is distance less than or equal to ϵ from K (discard cubes not intersecting K!) we have $K + \epsilon B \supseteq Q$. Since obviously also $Q + \epsilon B \supseteq K$, we have $d(Q, K) \leqslant \epsilon$.

(b) To construct P, cover K with finitely many ball of radius ϵ with centers in K. The convex hull of the set of all centers of these balls is a convex polytope $P \subseteq K$. Clearly $P + \epsilon B \supseteq K$. Also $K + \epsilon B \supseteq P$, so $d(P, K) \leqslant \epsilon$. ∎

Corollary 1 *Let K be a closed and bounded convex set in $\mathbf{R}^n$. Then there exists a sequence of convex polytopes $\{P_i\}$ with $P_i \subseteq K$, $i = 1, 2, \ldots$, and $P_i \to K$. Similarly there exists a sequence of convex polytopes $\{Q_i\}$ with $K \subseteq Q_i$, $i = 1, 2, \ldots$, and $Q_i \to K$.*

Theorem 32 *Let K be a closed and bounded convex set in $\mathbf{R}^n$, and suppose K contains a ball centered at the origin. Let $\lambda > 1$. Then there exists a convex polytope P with*

$$P \subseteq K \subseteq \lambda P.$$

Proof. Let $B = B(\mathbf{0}, 1)$. Choose $r > 0$ so small that $2rB \subset K$. Then there exists $\epsilon > 0$ such that

$$0 < \epsilon < (\lambda - 1)r, \text{ and } h(K, \mathbf{u}) > r + \epsilon, \text{ for all } \|\mathbf{u}\| = 1.$$

Let P be a polytope with $P \subseteq K$ and $d(P, K) \leqslant \epsilon$. Then

$$h(K, \mathbf{u}) < h(P, \mathbf{u}) + \epsilon, \text{ so } r < h(K, \mathbf{u}) - \epsilon < h(P, \mathbf{u}).$$

Hence

$$
\begin{aligned}
h(K, \mathbf{u}) &\leqslant h(P, \mathbf{u}) + \epsilon \\
&\leqslant h(P, \mathbf{u}) + (\lambda - 1)r \\
&< h(P, \mathbf{u}) + (\lambda - 1)h(P, \mathbf{u}) \\
&= \lambda h(P, \mathbf{u}) \\
&= h(\lambda P, \mathbf{u}).
\end{aligned}
$$

That is,

$$h(K, \mathbf{u}) < h(\lambda P, \mathbf{u}), \text{ for all } \|\mathbf{u}\| = 1.$$

Hence, by Exercise 10–6 (c), page 48,

$$K \subseteq \lambda P. \qquad \blacksquare$$

Let K be a closed and bounded convex set in $\mathbf{R}^n$. We denote the *volume* of K by $\mathrm{vol}(K)$. In $\mathbf{R}^2$ the *area* of K is denoted by $A(K)$ and the perimeter by $L(K)$.

Theorem 33 *If $K_i \to K$, then $\mathrm{vol}(K_i) \to \mathrm{vol}(K)$. (In other words, the volume of K is a continuous functional of K.)*

Proof in case $\mathrm{vol}(K) > 0$. In this case, assume $2rB \subseteq K$, $r > 0$, where $B = B(\mathbf{0}, 1)$. Let $0 < \epsilon < r$. We have then $h(K\mathbf{u}) > r$ for $\|\mathbf{u}\| = 1$. If $d(K_i, K) < \epsilon$, then

$$h(K_i, \mathbf{u}) < h(K, \mathbf{u}) + \epsilon < h(K, \mathbf{u}) + \frac{\epsilon}{r}h(K, \mathbf{u}) = h((1 + \frac{\epsilon}{r})K, \mathbf{u}).$$

Hence

$$K_i \subseteq (1 + \frac{\epsilon}{r})K,$$

So we have

$$\mathrm{vol}(K_i) \leqslant (1 + \frac{\epsilon}{r})^n \mathrm{vol}(K) \tag{7}$$

Next observe that

$$h(K, \mathbf{u}) < h(K_i, \mathbf{u}) + \epsilon < h(K_i, \mathbf{u}) + \frac{\epsilon}{r}h(K, \mathbf{u}).$$

Hence

$$(1 - \frac{\epsilon}{r})h(K, \mathbf{u}) < h(K_i, \mathbf{u}).$$

Since $1 - \frac{\epsilon}{r} > 0$, this implies that

$$h((1 - \frac{\epsilon}{r})K, \mathbf{u}) < h(K_i, \mathbf{u}).$$

Hence

$$(1 - \frac{\epsilon}{r})K \subseteq K_i,$$

So we have

$$\mathrm{vol}(K_i) \geqslant (1 - \frac{\epsilon}{r})^n \mathrm{vol}(K) \tag{8}$$

Equations 7 and 8 together give

$$(1 - \frac{\epsilon}{r})^n \mathrm{vol}(K) \leqslant \mathrm{vol}(K_i) \leqslant (1 + \frac{\epsilon}{r})^n \mathrm{vol}(K).$$

From this it follows that

$$\mathrm{vol}(K_i) \to \mathrm{vol}(K) \text{ as } d(K_i, K) \to 0. \qquad \blacksquare$$

Remark. In the case of $\mathbf{R}^2$ the preceding theorem tells us that if $K_i \to K$, then $A(K_i) \to A(K)$. Note that the proof could have been carried through in an analogous fashion for perimeter, rather than area. So for plane convex sets we also have

$$\text{if } K_i \to K, \text{ then } L(K_i) \to L(K).$$

This can also be proved directly from the formula

$$L(K) = \int_0^{2\pi} h(K, \theta) \, d\theta,$$

once the formula has been proved for all closed and bounded $K \subset \mathbf{R}^2$.
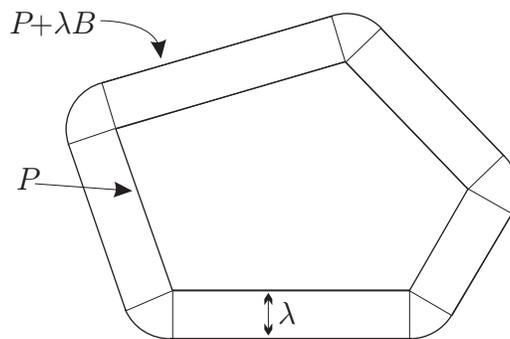
Figure 32: The outer parallel set of a polygon in $\mathbf{R}^2$.

16   The Outer Parallel Set of a Convex Set in $\mathbf{R}^2$

If P is a convex polygon in $\mathbf{R}^2$ (the conventional term for a convex polytope in $\mathbf{R}^2$!) then, recall from page 64, the outer parallel set at distance $\lambda > 0$ is $P + \lambda B$, where $B = B(\mathbf{0}, 1)$. One has for the area *Steiner's Formula*

$$A(P + \lambda B) = A(P) + \lambda L(P) + \pi \lambda^2.$$

Here $\lambda L(P)$ is the sum of the areas of rectangles and $\pi \lambda^2$ is the sum of the areas of sectors of circles (Figure 32).

The next theorem extends Steiner's Formula to any closed and bounded convex set in $\mathbf{R}^2$.

Theorem 34 *Let K be a closed and bounded convex set in $\mathbf{R}^2$. Then the area of the outer parallel set of K at distance λ satisfies*

$$A(K + \lambda B) = A(K) + \lambda L(K) + \pi \lambda^2.$$

Proof. Let $\{P_i\}$ be a sequence of convex polygons with $P_i \to K$. Then $A(P_i) \to A(K)$ and $L(P_i) \to L(K)$, so

$$A(P_i + \lambda B) = A(P_i) + \lambda L(P_i) + \pi \lambda^2 \to A(K) + \lambda L(K) + \pi \lambda^2.$$

But also

$$A(P_i + \lambda B) \to A(K + \lambda B), \text{ since } P_i + \lambda B \to K + \lambda B.$$

Thus

$$A(K + \lambda B) = A(K) + \lambda L(K) + \pi \lambda^2. \qquad \blacksquare$$

16.1   Exercises

16–1 The ellipse E in Example 9–E has support function

$$h(E, \mathbf{u}) = (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{\frac{1}{2}},$$

where $\mathbf{u} = (\cos \theta, \sin \theta)$. For convenience, we denote this function by

$$F(a, b, \theta) = (a^2 \cos^2 \theta + b^2 \sin^2 \theta)^{\frac{1}{2}}.$$

(a) Use your trigonometric powers to show that

$$(F(a, b, \theta) + F(b, a, \theta))^2 = a^2 + b^2 + 2\sqrt{a^2 b^2 + (a^2 - b^2)^2 \cos^2 \theta \sin^2 \theta}.$$

(b) From part (a) deduce that

$$a + b \leqslant F(a, b, \theta) + F(b, a, \theta) \leqslant \sqrt{2(a^2 + b^2)} \tag{9}$$

for all $0 \leqslant \theta \leqslant 2\pi$. Show also that if $0 \leqslant \theta \leqslant 2\pi$, with $\theta \neq 0$, $\frac{\pi}{2}$, $\frac{3\pi}{2}$, or 2π, and $a \neq b$, then

$$F(a, b, \theta) + F(b, a, \theta) > a + b.$$

[Hint for the proof of (9): For the lower bound $a + b$, observe that

$$(a^2 - b^2)^2 \cos^2 \theta \sin^2 \theta \geqslant 0.$$

For the upper bound $\sqrt{2(a^2 + b^2)}$, show that $(a^2 - b^2)^2 \cos^2 \theta \sin^2 \theta$ has a maximum value of $\frac{(a^2 - b^2)^2}{4}$.]

(c) Show that

$$\int_0^{2\pi} F(a, b, \theta) \, d\theta = \int_0^{2\pi} F(b, a, \theta) \, d\theta.$$

(d) Use the above to show that the perimeter of the ellipse E satisfies

$$\pi(a + b) \leqslant L(E) \leqslant \pi\sqrt{2(a^2 + b^2)},$$

and that, in fact, if $a \neq b$, then

$$\pi(a + b) < L(E) < \pi\sqrt{2(a^2 + b^2)}.$$

16–2 Let E be the plane convex set bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where $a > b > 0$. Steiner's formula, Theorem 34, for area gives us for the outer parallel set of E at distance λ,

$$A(E + \lambda B) = A(E) + \lambda L(E) + \pi \lambda^2.$$

Show that this parallel set <u>cannot</u> be an ellipse (if $a > b$ and $\lambda > 0$) by using the fact that $A(E) = \pi ab$ and that $L(E) > \pi(a+b)$ (from Exercise 16–1, part (d)). Note that if $E + \lambda B$ were an ellipse, then it would have semi-axes $a + \lambda$ and $b + \lambda$.

Remark. The perimeter of E has the form

$$L(E) = \int_0^{2\pi} \sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta} \, d\theta,$$

which turns out not to be expressible in terms of elementary functions. The integral is a so-called *elliptic integral.*

17   Inner Parallel Sets

Let K be a closed and bounded convex set in $\mathbf{R}^n$. In Section 14 the outer parallel set of K at distance λ, which we shall denote by K_λ, was defined as $K + \lambda B$, where $\lambda \geqslant 0$ and $B = B(\mathbf{0}, 1)$. Exercise 4–1 tells us that

$$K_\lambda = K + \lambda B = \{\mathbf{x} \in \mathbf{R}^n \ : \ d(\mathbf{x}, K) \leqslant \lambda\}.$$

This is easily shown to be the same as the locus of the centers of all closed balls of radius λ which intersect K. That is

$$K_\lambda = \{\mathbf{x} \in \mathbf{R}^n \ : \ B(\mathbf{x}, \lambda) \cap K \neq \emptyset\}.$$

We wish to extend the idea of outer parallel sets to inner parallel sets by defining $K_{-\lambda}$, $\lambda \geqslant 0$. For this, let r be the radius of the largest ball contained in K, so

$$r = \max\{\lambda \geqslant 0 \ : \ B(\mathbf{x}, \lambda) \subseteq K \text{ for some } \mathbf{x} \in \mathbf{R}^n\}.$$

We call r the *inradius* of K. Then, if $0 \leqslant \lambda \leqslant r$, the *inner parallel set of K at distance λ,* which we shall denote by $K_{-\lambda}$, is the subset of K defined by

$$K_{-\lambda} = \{\mathbf{x} \in K \ : \ B(\mathbf{x}, \lambda) \subseteq K\}.$$

K_0 is understood to be K itself.

A useful alternate characterization of inner parallel sets is as follows.

Theorem 35 *Let K be a closed and bounded convex set in $\mathbf{R}^n$ with $\mathbf{0} \in K$ and inradius $r > 0$. For each direction $\mathbf{u}$, with $\|\mathbf{u}\| = 1$, let $L_{\mathbf{u}}$ be the supporting hyperplane of K with equation*

$$h(K, \mathbf{u}) - \langle \mathbf{x}, \mathbf{u} \rangle = 0,$$

and let $H_{\mathbf{u}}$ be the halfspace determined by $L_{\mathbf{u}}$ containing K, with

$$H_{\mathbf{u}} = \{\mathbf{x} \in \mathbf{R}^n \ : \ h(K, \mathbf{u}) - \langle \mathbf{x}, \mathbf{u} \rangle \geqslant 0\}.$$
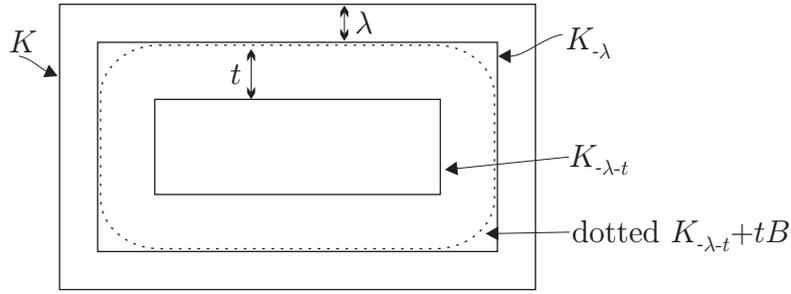
Figure 33: Theorem 37 $K_{-\lambda - t} + tB \subseteq K_{-\lambda}$.

Let $0 < \lambda \leqslant r$. For each direction $\mathbf{u}$, with $\|\mathbf{u}\| = 1$, let $H_{\mathbf{u}}^{\lambda}$ be the halfspace defined by

$$H_{\mathbf{u}}^{\lambda} = \{\mathbf{x} \in \mathbf{R}^n \ : \ h(K, \mathbf{u}) - \langle \mathbf{x}, \mathbf{u} \rangle \geqslant \lambda\}.$$

Then $K_{-\lambda}$ is the intersection of all the $H_{\mathbf{u}}^{\lambda}$, that is,

$$K_{-\lambda} = \bigcap_{\|\mathbf{u}\| = 1} H_{\mathbf{u}}^{\lambda}.$$

Remark. $H_{\mathbf{u}}^{\lambda}$ is simply the halfspace obtained by translating the supporting halfspace $H_{\mathbf{u}}$ "inward" through a distance λ.

The proof of the theorem is Exercise 17–1.

By translating the supporting halfspaces of K inward through distance λ, we obtain $K_{-\lambda}$. If we now translate the supporting halfspaces of $K_{-\lambda}$ inward through distance t, we obtain $(K_{-\lambda})_{-t}$ as their intersection, and we would expect this to be the same as $K_{-\lambda - t} = K_{-(\lambda + t)}$. This is the content of the next lemma.

Lemma 5 *Let K be a closed and bounded convex set in $\mathbf{R}^n$ with inradius $r > 0$. If $\lambda > 0$, $t > 0$ and $\lambda + t < r$, then*

$$K_{-\lambda - t} = (K_{-\lambda})_{-t}.$$

The proof is left as an exercise.

Theorem 36 *Let K be a closed and bounded convex set in $\mathbf{R}^n$ with inradius $r > 0$. Then, if $0 < t \leqslant r$, we have*

$$(K_{-t})_t \subseteq K,$$

that is, if $B = B(\mathbf{0}, 1)$, then $K_{-t} + tB \subseteq K$.

Proof. Since K_{-t} is the intersection of the halfspaces $H_{\mathbf{u}}^t$ (statement of Theorem 35), we have $h(K, \mathbf{u}) - \langle \mathbf{x}, \mathbf{u} \rangle \geqslant t$, for each $\mathbf{u}$, and all $\mathbf{x} \in K_{-t}$. That is

$$\langle \mathbf{x}, \mathbf{u} \rangle \leqslant h(K, \mathbf{u}) - t \text{ for all } \mathbf{x} \in K_{-t}.$$
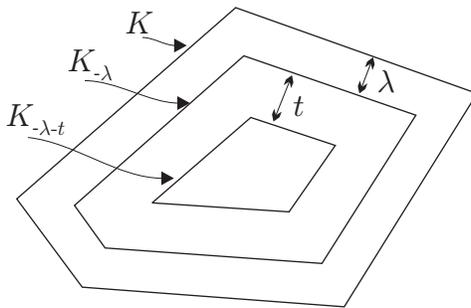
Figure 34: The inner parallel sets of a polygon K may have fewer edges than K.

Thus $h(K_{-t}, \mathbf{u}) = \max_{\mathbf{x} \in K_{-t}} \langle \mathbf{x}, \mathbf{u} \rangle \leqslant h(K, \mathbf{u}) - t$, so

$$h(K_{-t} + tB, \mathbf{u}) = h(K_{-t}, \mathbf{u}) + t \leqslant h(K, \mathbf{u}) \text{ for all } \mathbf{u}.$$

It follows from Exercise 10–6 (c), page 48, that $K_{-t} + tB \subseteq K$. ∎

Theorem 37 *Let K be a closed and bounded convex set in $\mathbf{R}^n$ with inradius $r > 0$. Then, given $t > 0$ and $t \leqslant \lambda \leqslant r + t$, we have*

$$K_{-\lambda} \supseteq K_{-\lambda-t} + tB,$$

(Figure 33).

The proof is left as an exercise.

17.1   Inner Parallel Sets of Convex Polygons

Recall that $A(K)$ and $L(K)$ represent the area and perimeter, respectively, of a plane convex set K (page 69).

Theorem 38 *Let K be a convex polygon in $\mathbf{R}^2$, with inradius $r > 0$. Then*

$$A(K) = \int_0^r L(K_{-\lambda}) \, d\lambda. \tag{10}$$

Proof. By Exercise 13–15 we see that $L(K_{-\lambda})$ is a decreasing function of λ on the interval $[0, r]$, so, by standard theorems of real analysis, $L(K_{-\lambda})$ is Riemann integrable on $[0, r]$. (In fact, it can be shown that $L(K_{-\lambda})$ is a continuous function of λ on $[0, r]$.)

Furthermore, the integral can be calculated by choosing subdivisions of $[0, r]$ of the form $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_{m-1} < \lambda_m = r$ with $\lambda_i = \frac{i}{m} r$, $i = 0, 1, \ldots, m$, and taking the limit of the corresponding Riemann sums as $m \to \infty$.

For simplicity of notation, let $A_i = A(K_{-\lambda_i})$ and $L_i = L(K_{-\lambda_i})$, $i = 0, 1, \ldots, m$. Observe that

$$A_{i-1} = A_i + (\lambda_i - \lambda_{i-1})L_i + \sum_j (\lambda_i - \lambda_{i-1})^2 \tan(\frac{\alpha_{ij}}{2}), \tag{11}$$
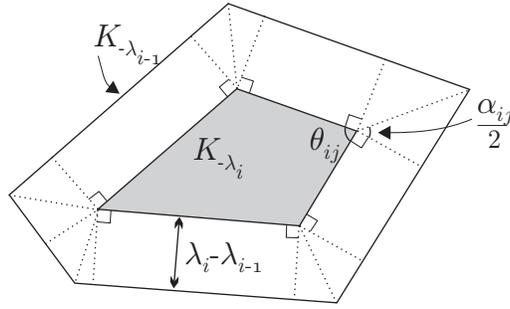
Figure 35: The region between parallel polygons partitioned as in (11).

where $\alpha_{ij} = \pi - \theta_{ij}$, and θ_{ij} ranges over the angles of $K_{-\lambda_i}$ (Figure 35). From (11) we have

$$|\sum_{i=1}^{m}(A_{i-1} - A_i) - \sum_{i=1}^{m} L_i(\lambda_i - \lambda_{i-1})| = |\sum_{i=1}^{m}\sum_{j}(\lambda_i - \lambda_{i-1})^2 \tan(\frac{\alpha_{ij}}{2})|.$$

But

$$\sum_{i=1}^{m}(A_{i-1} - A_i) = A_0 - A_m = A(K) - A(K_{-r}) = A(K),$$

since $A(K_{-r}) = 0$ (Exercise 17–5), so we have

$$|A(K) - \sum_{i=1}^{m} L_i(\lambda_i - \lambda_{i-1})| = |\sum_{i=1}^{m}\sum_{j}(\lambda_i - \lambda_{i-1})^2 \tan(\frac{\alpha_{ij}}{2})|.$$

Next we observe that the angles of $K_{-\lambda_i}$ are formed by pairs of lines parallel to nonparallel sides of K (not necessarily adjacent sides!) (Figure 34) and there are only finitely many such sides. Thus there exists a θ_0 such that $\theta_{ij} > \theta_0 > 0$, for all ij. So $\alpha_{ij} < \pi - \theta_0$ and $\alpha_{ij}/2 < \frac{\pi}{2} - \theta_0/2 < \frac{\pi}{2}$. It follows that there is a constant M such that

$$\sum_{j} |\tan(\frac{\alpha_{ij}}{2})| \leqslant M$$

for all $i = 0, 1, \ldots, m$. So we have

$$|A(K) - \sum_{i=1}^{m} L_i(\lambda_i - \lambda_{i-1})| \leqslant M \sum_{i=1}^{m}(\lambda_i - \lambda_{i-1})^2.$$

But with our choice of subdivisions of $[0, r]$, we have $\lambda_i - \lambda_{i-1} = \frac{r}{m}$, so

$$|A(K) - \sum_{i=1}^{m} L_i(\lambda_i - \lambda_{i-1})| \leqslant M \frac{r^2}{m}.$$

The righthand side tends to 0 as $m \to \infty$, so the Riemann sums approach $A(K)$ as $m \to \infty$, proving (10).   ∎

Remark. It can be shown that formula (10) for area is valid for any closed and bounded plane convex set, and we shall apply it with this generality, particularly in a proof of the Isoperimetric Inequality in the next section.

17.2   Exercises

17–1 Prove Theorem 35 as follows.
   (a) Show that $K_{-\lambda} \subseteq \bigcap_{\|\mathbf{u}\|=1} H_{\mathbf{u}}^{\lambda}$. [Hint: If $\mathbf{k} \in K_{-\lambda}$, then $\mathbf{k} + \lambda\mathbf{u} \in K$ for all $\mathbf{u}$ with $\|\mathbf{u}\| = 1$.]
   (b) Next show that $K_{-\lambda} \supseteq \bigcap_{\|\mathbf{u}\|=1} H_{\mathbf{u}}^{\lambda}$. [Hint: If $\mathbf{k} \in \bigcap_{\|\mathbf{u}\|=1} H_{\mathbf{u}}^{\lambda}$ use Exercise 10–6 (c) to show that $B(\mathbf{x}, \lambda) \subseteq K$.]

17–2 Prove Lemma 5.

17–3 Prove Theorem 37. [Hint: In Theorem 36 substitute $K_{-\lambda}$ for K and use the fact that $K_{-\lambda-t} = (K_{-\lambda})_{-t}$.]

17–4 Let F_1, F_2, ..., F_m be the facets of a convex polytope K in $\mathbf{R}^n$. Then F_i is contained in a hyperplane L_i with equation $h(K, \mathbf{u}_i) - \langle \mathbf{x}, \mathbf{u}_i \rangle = 0$, where $\|\mathbf{u}_i\| = 1$, $i = 1, 2, \ldots, m$, and K is the intersection of the halfspaces

$$H_i = \{\mathbf{x} \in \mathbf{R}^n \;:\; h(K, \mathbf{u}_i) - \langle \mathbf{x}, \mathbf{u}_i \rangle \geqslant 0\}, \; i = 1, 2, \ldots, m.$$

Let r be the inradius of K and $0 \leqslant \lambda \leqslant r$. Show that $K_{-\lambda}$ is the intersection of the halfspaces

$$H_i^{\lambda} = \{\mathbf{x} \in \mathbf{R}^n \;:\; h(K, \mathbf{u}_i) - \langle \mathbf{x}, \mathbf{u}_i \rangle \geqslant \lambda\}, \; i = 1, 2, \ldots, m,$$

and therefore $K_{-\lambda}$ is a convex polytope for $0 \leqslant \lambda \leqslant r$.

17–5 Let K be a closed and bounded convex set in $\mathbf{R}^n$ with inradius $r > 0$. Prove that K_{-r} has dimension less that n, so $\mathrm{vol}(K_{-r}) = 0$. [Hint: Show that if K_{-r} contains a ball $B(\mathbf{x}, \epsilon)$, $\epsilon > 0$, then K contains $B(\mathbf{x}, \epsilon) + rB(\mathbf{0}, 1) = B(\mathbf{x}, r + \epsilon)$.]

18   The Isoperimetric Theorem

The classic Isoperimetric Theorem asserts that of all simple closed curves (closed curves without self-intersection) in $\mathbf{R}^2$ having the same perimeter, the circle encloses that largest area. In this section we consider this theorem for convex sets.

   We can reformulate the result in the form of an inequality. Let K be a closed and bounded convex set in $\mathbf{R}^2$, having perimeter $L(K)$ and area $A(K)$. A circle of the same perimeter has radius $\rho = \frac{L(K)}{2\pi}$ and area $\pi\rho^2 = \pi(\frac{L(K)}{2\pi})^2 = \frac{L(K)^2}{4\pi}$. The fact that the area of the circle is at least that of K can then be expressed as the following inequality, called the *Isoperimetric Inequality*:

$$\frac{L(K)^2}{4\pi} \geqslant A(K).$$

This can equivalently be stated in the following form (keep in mind that we are restricting ourselves to convex sets, although the result holds for a wider class of regions.)

Theorem 39 *Let K be a closed and bounded convex set in $\mathbf{R}^2$ with perimeter L and area A. Then*

$$L^2 - 4\pi A \geqslant 0,$$

and equality holds if and only if K is a circular disk.

Remark. The last clause in the theorem tells us that the circle is the <u>unique</u> curve that maximizes enclosed area among all curves of the same perimeter.

Proof. Let $r > 0$ be the inradius of K, and $B = B(\mathbf{0}, 1)$. We have

$$K \supseteq K_{-\lambda} + \lambda B, \ \text{ for } 0 \leqslant \lambda \leqslant r.$$

Therefore (using Exercises 13–10 and 13–15)

$$L(K) \geqslant L(K_{-\lambda} + \lambda B) \geqslant L(K_{-\lambda}) + L(\lambda B) = L(K_{-\lambda}) + 2\pi\lambda.$$

Thus

$$L(K_{-\lambda}) \leqslant L(K) - 2\pi\lambda, \ \text{ for } 0 \leqslant \lambda \leqslant r.$$

Therefore

$$A = A(K) = \int_0^r L(K_{-\lambda}) \, d\lambda \leqslant rL(K) - \pi r^2 = rL - \pi r^2, \ \text{ or}$$

$$-A + rL - \pi r^2 \geqslant 0.$$

This last inequality is known as *Bonnesen's inequality*, and it immediately implies the Isoperimetric inequality, because

$$L^2 - 4\pi A = (L - 2\pi r)^2 + 4\pi(-A + rL - \pi r^2) \geqslant 0.$$

Thus we have $L^2 - 4\pi A \geqslant 0$, and $L^2 - 4\pi A = 0$ only if both

$$L = 2\pi r \ \text{ and } \ -A + rL - \pi r^2 = 0.$$

In particular, if $L = 2\pi r$, then K coincides with its inscribed circle (Exercise 13–15).   ∎

18.1   Exercises

18–1 (a) Let P be a convex polygon circumscribed about a circle of radius r (so each edge of P is tangent to the circle). If P has perimeter L and area A, show that

$$A = \frac{1}{2}rL.$$

(b) Use (a) to directly check that P satisfies the Isoperimetric inequality.

18–2 Suppose K is such that
$$-A + rL - \pi r^2 = 0.$$

(a) From the proof of the Isoperimetric inequality, deduce that

$$L(K_{-\lambda}) = L(K) - 2\pi\lambda, \quad \text{for } 0 \leqslant \lambda \leqslant r,$$

and in particular,
$$L(K) = L(K_{-r}) + 2\pi r = L(K_{-r} + rB).$$

(b) From (a) deduce that
$$K = K_{-r} + rB.$$

(c) Since K_{-r} is either a point or a line segment (Exercise 17–5), deduce that if $-A(K) + rL(K) - \pi r^2 = 0$, then either K is either a circular disk or the outer parallel set of a line segment.

18–3 Let E be an ellipse with semi-axes a and b. From Exercise 16–1 (d) we have

$$L(E) \geqslant \pi(a + b).$$

Use this to show that if $A(E)$ is the area of E, then

$$L(E)^2 - 4\pi A(E) \geqslant \pi^2(a - b)^2 \geqslant 0,$$

with equality if and only if E is a circle.

18–4 If R is a rectangle with sides a and b, then

$$L(R) = 2(a + b) \text{ and } A(R) = ab.$$

The Isoperimetric inequality tells us that

$$L(R)^2 \geqslant 4\pi A(R).$$

Show that we, in fact, have a stronger inequality

$$L(R)^2 \geqslant 16A(R),$$

with equality if and only if R is a square.

18–5 Bonnesen's inequality for a plane convex set of area A, perimeter L and inradius r is

$$-A + rL - \pi r^2 \geqslant 0.$$

Show that for a rectangle (previous exercise) we have

$$-A + rL - 4r^2 \geqslant 0.$$

18–6 (a) Let P be any plane convex quadrilateral with consecutive sides of lengths, a, b, c, d, so $L(P) = a + b + c + d$. Let α, β, γ, δ be the consecutive angles of P, with α the angle between the sides of lengths a and b, β the angle between the sides of lengths b and c, and so forth. Show that

$$4A(P) = ab \sin \alpha + bc \sin \beta + cd \sin \gamma + da \sin \delta,$$

and from this deduce

$$4A(P) \leqslant (a + c)(b + d), \tag{12}$$

with equality if and only if P is a rectangle.

(b) From part (a) and the Arithmetic-Geometric Mean inequality, derive

$$L(P)^2 \geqslant 16A(P),$$

with equality if and only if P is a square.

Remark. Inequality (12) is the *Isoperimetric inequality for quadrilaterals.* It is equivalent to the statement that among all plane quadrilaterals of the same perimeter, the square has the largest area. The above argument was made for convex quadrilaterals, but the result holds for all plane quadrilaterals, as can be seen by applying the inequality to the convex hull of any plane quadrilateral.

18–7 The *Isoperimetric inequality for n–gons* (which we do not prove here) asserts that any plane n–gon with area A and perimeter L satisfies

$$L^2 \geqslant 4n \tan(\frac{\pi}{n}) A.$$

Explain why this is equivalent to the statement that a regular n–gon encloses area at least as large as any n–gon of the same perimeter.

18–8 Let K be a closed and bounded plane convex set with inradius $r > 0$. Define

$$f(\lambda) = (L(K_{-\lambda}))^2, \ \text{ for } 0 \leqslant \lambda \leqslant r.$$

For the purposes of this exercise, assume f is differentiable at all but a countable number of points in $[0, r]$.

(a) Using Theorem 37, show that

$$f(\lambda) \geqslant f(\lambda + t) + 4\pi t(L(K_{-\lambda-t})),$$

and deduce from this that for those values of λ such that the derivative $f'(\lambda)$ exists, we have

$$-f'(\lambda) \geqslant 4\pi L(K_{-\lambda}).$$

(b) Deduce from part (a) that

$$L(K)^2 - (L(K_{-r}))^2 \geqslant 4\pi A(K).$$

That is, we have a strengthening of the Isoperimetric inequality in the form

$$L(K)^2 - 4\pi A(K) \geqslant (L(K_{-r}))^2 \geqslant 0.$$
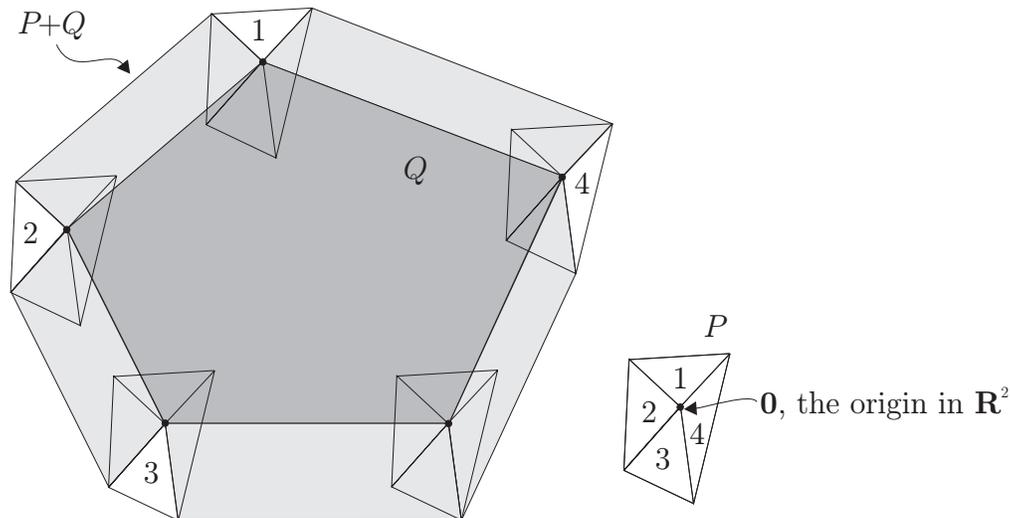
Figure 36: Partitioning P and $P + Q$ to illustrate $A(P, Q)$.

19   Mixed Areas in **R**^{2}

In this section we define a functional called the mixed area of two plane convex sets which assigns to K_1 and K_2 a real number $A(K_1, K_2)$. It will turn out that for any planar convex set K, $A(K, K) = A(K)$ and $A(K, B) = \frac{L(K)}{2}$ when B is the circular disk of radius 1.

There are several avenues open to us for introducing the concept of mixed area. We choose to start with the case where P and Q are convex polygons. The properties we derive will be extended to general closed and bounded plane convex sets using approximation by polygons.

In Figure 36 we see that the polygon $P + Q$ can be partitioned into pieces:

- Q (dark shading),

- parallelograms (light shading), and

- a dissection of P (polygons labeled 1, 2, 3 and 4 in the figure),

giving
$$A(P + Q) = A(Q) + (\text{sum of area of parallelograms}) + A(P).$$

The sum of the area of the parallelograms is

$$\sum_{i=1}^{n} h(P, \theta_i)\ell(Q, \theta_i),$$

where Q has n edges, the outer normal to the i-th edge is $(\cos\theta_i, \sin\theta_i)$, $\ell(Q, \theta_i)$ is the length of the i-th edge of Q, and $h(P, \theta_i)$ is the support function of P in direction $(\cos\theta_i, \sin\theta_i)$, $i = 1, 2, \ldots, n$.

We define the *mixed area of two polygons* $A(P, Q)$ to be

$$A(P, Q) \stackrel{\text{def}}{=} \sum_{i=1}^{n} h(P, \theta_i) \ell(Q, \theta_i). \tag{13}$$

With this definition we have

$$A(P + Q) = A(P) + 2A(P, Q) + A(Q).$$

19.1 Extension of Mixed Areas to General Convex Sets

Let K_1 and K_2 be closed and bounded plane convex sets, not necessarily polygons, and choose any sequences of convex polygons $\{P_n\}$ and $\{Q_n\}$, such that

$$P_n \to K_1 \text{ and } Q_n \to K_2 \text{ as } n \to \infty.$$

The Corollary to Theorem 31, page 68, guarantees the existence of such sequences. From Exercise 14–5, we also have $(P_n + Q_n) \to (K_1 + K_2)$. Then, by the definition above,

$$A(P_n, Q_n) = \sum_{i=1}^{m_n} h(P_n, \theta_i) \ell(Q_n, \theta_i),$$

where polygon Q_n has m_n edges. We now define the *mixed area* of K_1 and K_2 by

$$A(K_1, K_2) \stackrel{\text{def}}{=} \frac{1}{2}(A(K_1 + K_1) - A(K_1) - A(K_2)). \tag{14}$$

Since

$$A(P_n) \to A(K_1), \ A(Q_n) \to A(K_2), \ \text{ and } A(P_n + Q_n) \to A(K_1 + K_2) \text{ as } n \to \infty,$$

by Theorem 33, page 70, we have

$$A(P_n, Q_n) = \frac{1}{2}(A(P_n + Q_n) - A(P_n) - A(Q_n)) \to \frac{1}{2}(A(K_1 + K_1) - A(K_1) - A(K_2)) = A(K_1, K_2).$$

19.2 Exercises

Prove the properties in Exercises 19–1 to 19–8 for all closed and bounded plane convex sets. Each property may be established for polygons using (13) and then extending it to closed and bounded plane convex sets using approximation, as in the previous section, or using (14), which applies to all sets.

19–1 $A(K_1, K_2) = A(K_2, K_1)$ (symmetry).

19–2 $A(K, K) = A(K)$.
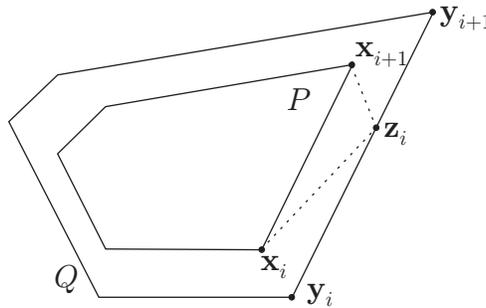
Figure 37: P and Q are convex polygons with parallel edges (Exercise 19–9).

19–3 $A(\mathbf{a} + K_1, \mathbf{b} + K_2) = A(K_1, K_2)$ (translation invariance).

19–4 $A(K, B) = \frac{1}{2} L(K)$, where $B = B(\mathbf{0}, 1)$.

19–5 $A(\lambda K_1, K_2) = \lambda A(K_1, K_2) = A(K_1, \lambda K_2)$, for all $\lambda > 0$ (homogeneous of degree 1 in each variable).

19–6 If $K_1 \subseteq K_2$, then $A(K_1, K_3) \leqslant A(K_2, K_3)$, and similarly in the second variable (monotonicity).

19–7 $A(K_1 + K_2, K_3) = A(K_1, K_2) + A(K_2, K_3)$.

19–8 $A(\lambda_1 K_1 + \lambda_2 K_2) = \lambda_1^2 A(K_1) + 2\lambda_1 \lambda_2 A(K_1, K_2) + \lambda_2^2 A(K_2)$.

19–9 Let P and Q be convex n–gons in $\mathbf{R}^2$, with $P \subset Q$ and the sides of P parallel to corresponding sides of Q. If $\overline{\mathbf{x}_i \mathbf{x}_{i+1}}$ is a side of P and $\overline{\mathbf{y}_i \mathbf{y}_{i+1}}$ the corresponding parallel side of Q, choose any point $\mathbf{z}_i \in \overline{\mathbf{y}_i \mathbf{y}_{i+1}}$, with $\mathbf{z}_n \in \overline{\mathbf{y}_n \mathbf{y}_1}$ (Figure 37). Let R be the (nonconvex) polygon with vertices

$$\mathbf{x}_1, \ \mathbf{z}_1, \ \mathbf{x}_2, \ \mathbf{z}_2, \ \mathbf{x}_3, \ \mathbf{z}_3, \ldots, \mathbf{x}_{n-1}, \ \mathbf{z}_{n-1}, \ \mathbf{x}_n, \ \mathbf{z}_n, \ \mathbf{x}_1.$$

Show that the area enclosed by R is $A(P, Q)$.

19–10 Suppose $\{K_n\}$ and $\{L_n\}$ are sequences such that $K_n \to K$ as $n \to \infty$ and $L_n \to L$ as $n \to \infty$. Prove that
$$A(K_n, L_n) \to A(K, L) \text{ as } n \to \infty.$$
[Hint: $2A(K_n, L_n) = A(K_n + L_n) - A(K_n) - A(L_n)$. Recall Exercise 14–5.]

19–11 Suppose $\{K_n\}$ is a sequence such that $K_n \to K$ as $n \to \infty$. Use Exercise 19–10 to prove that $L(K_n) \to L(K)$ as $n \to \infty$. [Hint: See Exercise 19–4.]

19–12 Of all plane convex sets of constant width 1, explain why a circular disk of diameter 1 has <u>maximum</u> area. [Hint: Recall Barbier's Theorem, Exercise 13–11 (b).]

19–13 The previous exercise might bring to mind the question of what figure has <u>minimum</u> area among all plane convex sets of constant width 1. The answer is the Reuleaux triangle of constant width 1, page 37. This result is called the *Blaschke-Lebesgue Theorem*, whose proof is developed in this exercise. (However, we do not prove a major clause of this theorem, namely, that the Reuleaux triangle is the <u>unique</u> figure that minimizes area among all plane convex sets of the same constant width.)

So, let K be a convex set of constant width 1 in $\mathbf{R}^2$. Let H be a regular hexagon of side length $\frac{1}{\sqrt{3}}$ circumscribed about K, as constructed in our proof of Pál's Theorem, Theorem 24, page 58. Assume K and H are positioned so that $\mathbf{0}$ is the center H, hence $-H = H$.

(a) Check that $A(H) = \frac{\sqrt{3}}{2}$.

(b) Explain why $A(K, -K) \leqslant A(H, H) = A(H) = \frac{\sqrt{3}}{2}$. [Hint: See Exercise 19–6.]

(c) Show that $A(K + (-K)) = \pi$, and deduce from this that

$$A(K) \geqslant \frac{\pi - \sqrt{3}}{2}.$$

Why does this show that K has area at least that of a Reuleaux triangle of constant width 1?

20 Minkowski's Inequality for Plane Convex Sets

Minkowski's inequality for convex sets in $\mathbf{R}^2$ (not the algebraic inequality derived in Exercise 2–6) gives a generalization of the Isoperimetric inequality.

Theorem 40 (Minkowski's Inequality for Planar Convex Sets) *Let K_1 and K_2 be closed and bounded convex sets in* $\mathbf{R}^2$. *Then*

$$A(K_1, K_2)^2 \geqslant A(K_1)A(K_2), \tag{15}$$

and equality holds if and only if K_1 is a homothet to K_2, that is, K_2 is a translate of λK_1 for some $\lambda > 0$.

Remark. If $K_1 = K$ and $K_2 = B = B(\mathbf{0}, 1)$, then $A(K_1, K_2) = A(K, B) = L(K)/2$, $A(K_1) = A(K)$, and $A(K_2) = \pi$. Then Minkowski's inequality reduces to

$$L(K)^2 \geqslant 4\pi A(K),$$

with equality if and only if K is a circular disk.

We give a proof of the inequality using the method of inner parallel bodies, but do not treat the case of equality.

Proof part (i). Assume K_1 and K_2 have the same inradius $r > 0$. For notational conve-
nience, denote $K_1 = K$ and $K_2 = M$ provisionally. Let $B = B(\mathbf{0}, 1)$. For $\lambda > 0$, $t > 0$,
$\lambda + t < r$, recall that

$$K_{-(\lambda+t)} + tB \subseteq K_{-\lambda} \text{ and } M_{-(\lambda+t)} + tB \subseteq M_{-\lambda}$$

(Section 17). By the monotonicity of mixed area (Exercise 19–6), applied to each variable,
we have

$$A(K_{-(\lambda+t)} + tB, M_{-(\lambda+t)} + tB) \leqslant A(K_{-\lambda}, M_{-\lambda}).$$

This yields

$$A(K_{-\lambda}, M_{-\lambda}) \geqslant A(K_{-(\lambda+t)}, M_{-(\lambda+t)}) + t(A(K_{-(\lambda+t)}, B) + A(M_{-(\lambda+t)}, B)) + \pi t^2.$$

But $A(K, B) = L(K)/2$ and $\pi t^2 > 0$, so we obtain

$$A(K_{-\lambda}, M_{-\lambda}) - A(K_{-(\lambda+t)}, M_{-(\lambda+t)}) \geqslant \frac{t}{2}(L(K_{-(\lambda+t)}) + L(M_{-(\lambda+t)})).$$

Now let $0 = \lambda_0 < \lambda_2 < \cdots < \lambda_{m-1} < \lambda_m = r$ be a partition of the interval $[0, r]$, Apply the
last inequality with $\lambda = \lambda_{i-1}$ and $t = \lambda_i - \lambda_{i-1}$ to get

$$A(K_{-\lambda_{i-1}}, M_{-\lambda_{i-1}}) - A(K_{-\lambda_i}, M_{-\lambda_i}) \geqslant \frac{1}{2}(L(K_{-\lambda_i}) + L(M_{-\lambda_i}))(\lambda_i - \lambda_{i-1}),$$

$i = 1, 2, \ldots, m$. Sum this for $i = 1, 2, \ldots, m$ and note that the lefthand side is a telescoping
sum reducing to

$$A(K_0, M_0) - A(K_{-r}, M_{-r}) = A(K, M) - A(K_{-r}, M_{-r}).$$

The righthand side is a Riemann sum, and we have

$$A(K, M) - A(K_{-r}, M_{-r}) \geqslant \frac{1}{2}\sum_{i=1}^{m}(L(K_{-\lambda_i}) + L(M_{-\lambda_i}))(\lambda_i - \lambda_{i-1}).$$

As we let the mesh of the partition approach 0, the righthand side approaches the Riemann
integral, so we have

$$A(K, M) - A(K_{-r}, M_{-r}) \geqslant \frac{1}{2}\int_0^r L(K_{-\lambda}) + L(M_{-\lambda}) \, d\lambda.$$

Using (10) and the Remark on page 76, we have

$$A(K, M) \geqslant A(K_{-r}, M_{-r}) + \frac{1}{2}(A(K) + A(M)).$$

Reverting back to our notation K_1 and K_2, and using $A(K_{-r}, M_{-r}) \geqslant 0$, we have

$$A(K_1, K_2) \geqslant \frac{1}{2}(A(K_1) + A(K_2)), \tag{16}$$
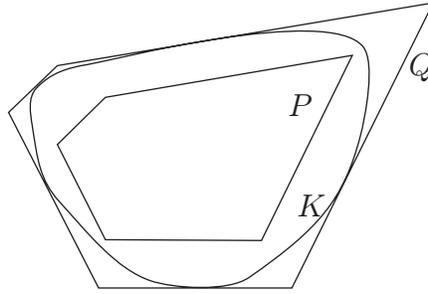
Figure 38: Q is circumscribed about K, $P \subset K$ (Exercise 20–1).

when K_1 and K_2 have the same inradius.

Part (ii). Given any closed and bounded plane convex sets K_1 and K_2 with inradii r_1 and r_2 respectively, $(\frac{1}{r_1})K_1$ and $(\frac{1}{r_2})K_2$ have the same inradius $r = 1$, so (16) gives

$$A(\frac{1}{r_1}K_1, \frac{1}{r_2}K_2) \geqslant \frac{1}{2}(A(\frac{1}{r_1}K_1) + A(\frac{1}{r_2}K_2)).$$

Using homogeneity (Exercise 19–5), this implies

$$\frac{1}{r_1 r_2}A(K_1, K_2) \geqslant \frac{1}{2}(\frac{1}{r_1^2}A(K_1) + \frac{1}{r_2^2}A(K_2)).$$

From this we get

$$
\begin{aligned}
A(K_1, K_2) \quad &\geqslant \quad \frac{1}{2}(\frac{r_2}{r_1}A(K_1) + \frac{r_1}{r_2}A(K_2)) \\
&= \quad \frac{1}{2}\left(\sqrt{\frac{r_2}{r_1}A(K_1)} - \sqrt{\frac{r_1}{r_2}A(K_2)}\right)^2 + \sqrt{A(K_1)A(K_2)},
\end{aligned}
$$

giving us the required Minkowski inequality (15). ∎

20.1 Exercises

20–1 Let K be a closed and bounded 2–dimensional convex set in $\mathbf{R}^2$. Suppose P is a convex n–gon with $P \subset K$, and Q a convex n–gon circumscribed about K with sides parallel to those of P (Figure 38). Show that

$$A(K) \geqslant \sqrt{A(P)A(Q)}.$$

[Hint: Use Exercise 19–9 to show that $A(P, Q) \leqslant A(K)$; then apply Minkowski's inequality for convex sets.]

20–2 Suppose K_1 and K_2 are closed and bounded convex sets in $\mathbf{R}^2$, and

$$K = \lambda_1 K_1 + \lambda_2 K_2,$$

where $\lambda_1 \geqslant 0$ and $\lambda_2 \geqslant 0$. Show that

$$\sqrt{A(K)} \geqslant \lambda_1 \sqrt{A(K_1)} + \lambda_2 \sqrt{A(K_2)},$$

with equality if and only if K_1 is homothetic to K_2. [Hint: Use Exercise 19–8.]

Remark. This is known as the *Brunn-Minkowski theorem* for plane convex sets. It is equivalent to Minkowski's inequality for plane convex sets.

20–3 Suppose P is a plane convex n–gon, and Q is a convex n–gon of the same perimeter with sides parallel to corresponding to sides of P and circumscribed about a circle. Show that

$$A(Q) \geqslant A(P).$$

[Hint: If Q is circumscribed about a circle of radius r, show that

$$A(Q) = rL(Q)/2 = rL(P)/2.$$

Also show that $A(P, Q) = rL(P)/2$. Apply Minkowski's inequality.]

20–4 From the previous exercise deduce that among all rectangles of the same perimeter, the square has largest area. (This is, of course, a much weaker result than Exercise 18–6 (b), page 80.)

References

Akerberg, B. 1963, A proof of the Arithmetic-Geometric Mean Inequality, *Amer. Math. Monthly* **70**, 997–998.

Bonnesen, T. & Fenchel, W. 1934, *Theorie der konvexen Körper.* Springer, Berlin. Reprint: Chelsea Publ. Co., New York, 1948. English translation: BCS Associates, Moscow, Idaho, 1987.

Brass, P., Moser, W. & Pach, J. 2005, *Research in Discrete Geometry.* Springer, New York.

Collier, J. B. 1976, On the facial structure of a convex body, *Proc. Amer. Math. Soc.* **61**, no. 2, 367–370.

Coxeter, H. S. M. 1948, *Regular Polytopes.* Methuen, London. Reprinted: Dover, New York, 1973.

Croft, H. T., Falconer, K. J. & Guy, R. K. 1991, *Unsolved Problems in Geometry.* Springer-Verlag, New York.

Eckhoff, J. 1993, Helly, Radon, and Carathéodory type theorems. In: *Handbook of Convex Geometry*, vol. A (eds P. M. Gruber & J. M. Wills), North-Holland, Amsterdam, 389–448.

Eggleston, H. G. 1969, *Convexity.* Cambridge Univ. Press.

Gardner, M. 1975, *Mathematical Carnival.* Knopf, New York.

Grünbaum, B. 1959, On intersections of similar sets, *Portugal. Math.* **18**, 155–164.
———— 2003, *Convex Polytopes*, second edition. Springer.

Hardy, G. H., Littlewood, J. E. & Pólya, G., 1952, *Inequalities.* Cambridge Univ. Press.

Hilbert, D. & Cohn-Vossen, S. 1952, *Geometry and the Imagination.* Chelsea.

Kahn, J. & Kalai, G. 1993, A counterexample to Borsuk's conjecture, *Bull. Amer. Math. Soc.*, New Ser. **29**, 60–62.

Karasev, R. N. 2000, Transversals for families of a two-dimensional convex compact set, *Discrete Comput. Geom.* **24**, 345–353.

Klee V. 1958, Extremal Structure of Convex Sets II, *Math. Z.* **69**, 90-104.

Pál, J. 1920, Über ein elementares Variationsproblem, *Danske Vid. Selsk. Math. Fys. Medd.* **3**, no. 2, 35p.

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

Author Index

Symbols

Index