Notes on Convex Sets, Polytopes, Polyhedra, Combinatorial Topology, Voronoi Diagrams and Delaunay Triangulations

Jean Gallier
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@cis.upenn.edu

June 30, 2009

Notes on Convex Sets, Polytopes, Polyhedra, Combinatorial Topology, Voronoi Diagrams and Delaunay Triangulations

Jean Gallier

Abstract: Some basic mathematical tools such as convex sets, polytopes and combinatorial topology, are used quite heavily in applied fields such as geometric modeling, meshing, computer vision, medical imaging and robotics. This report may be viewed as a tutorial and a set of notes on convex sets, polytopes, polyhedra, combinatorial topology, Voronoi Diagrams and Delaunay Triangulations. It is intended for a broad audience of mathematically inclined readers.

One of my (selfish!) motivations in writing these notes was to understand the concept of *shelling* and how it is used to prove the famous Euler-Poincaré formula (Poincaré, 1899) and the more recent *Upper Bound Theorem* (McMullen, 1970) for polytopes. Another of my motivations was to give a "correct" account of Delaunay triangulations and Voronoi diagrams in terms of (direct and inverse) stereographic projections onto a sphere and prove rigorously that the projective map that sends the (projective) sphere to the (projective) paraboloid works correctly, that is, maps the Delaunay triangulation and Voronoi diagram w.r.t. the lifting onto the sphere to the Delaunay diagram and Voronoi diagrams w.r.t. the traditional lifting onto the paraboloid. Here, the problem is that this map is only well defined (total) in projective space and we are forced to define the notion of convex polyhedron in projective space.

It turns out that in order to achieve (even partially) the above goals, I found that it was necessary to include quite a bit of background material on convex sets, polytopes, polyhedra and projective spaces. I have included a rather thorough treatment of the equivalence of $\mathcal{V}$-polytopes and $\mathcal{H}$-polytopes and also of the equivalence of $\mathcal{V}$-polyhedra and $\mathcal{H}$-polyhedra, which is a bit harder. In particular, the *Fourier-Motzkin elimination* method (a version of Gaussian elimination for inequalities) is discussed in some detail. I also had to include some material on projective spaces, projective maps and polar duality w.r.t. a nondegenerate quadric in order to define a suitable notion of "projective polyhedron" based on cones. To the best of our knowledge, this notion of projective polyhedron is new. We also believe that some of our proofs establishing the equivalence of $\mathcal{V}$-polyhedra and $\mathcal{H}$-polyhedra are new.

Key-words: Convex sets, polytopes, polyhedra, shellings, combinatorial topology, Voronoi diagrams, Delaunay triangulations.

Contents

Chapter 1

Introduction

1.1 Motivations and Goals

For the past eight years or so I have been teaching a graduate course whose main goal is to expose students to some fundamental concepts of geometry, keeping in mind their applications to geometric modeling, meshing, computer vision, medical imaging, robotics, *etc.* The audience has been primarily computer science students but a fair number of mathematics students and also students from other engineering disciplines (such as Electrical, Systems, Mechanical and Bioengineering) have been attending my classes. In the past three years, I have been focusing more on convexity, polytopes and combinatorial topology, as concepts and tools from these areas have been used increasingly in meshing and also in computational biology and medical imaging. One of my (selfish!) motivations was to understand the concept of *shelling* and how it is used to prove the famous Euler-Poincaré formula (Poincaré, 1899) and the more recent *Upper Bound Theorem* (McMullen, 1970) for polytopes. Another of my motivations was to give a "correct" account of Delaunay triangulations and Voronoi diagrams in terms of (direct and inverse) stereographic projections onto a sphere and prove rigorously that the projective map that sends the (projective) sphere to the (projective) paraboloid works correctly, that is, maps the Delaunay triangulation and Voronoi diagram w.r.t. the lifting onto the sphere to the Delaunay triangulation and Voronoi diagram w.r.t. the lifting onto the paraboloid. Moreover, the projections of these polyhedra onto the hyperplane $x_{d+1} = 0$, from the sphere or from the paraboloid, are identical. Here, the problem is that this map is only well defined (total) in projective space and we are forced to define the notion of convex polyhedron in projective space.

It turns out that in order to achieve (even partially) the above goals, I found that it was necessary to include quite a bit of background material on convex sets, polytopes, polyhedra and projective spaces. I have included a rather thorough treatment of the equivalence of $\mathcal{V}$-polytopes and $\mathcal{H}$-polytopes and also of the equivalence of $\mathcal{V}$-polyhedra and $\mathcal{H}$-polyhedra, which is a bit harder. In particular, the *Fourier-Motzkin elimination* method (a version of Gaussian elimination for inequalities) is discussed in some detail. I also had to include some material on projective spaces, projective maps and polar duality w.r.t. a nondegenerate quadric, in order to define a suitable notion of "projective polyhedron" based on cones. This notion turned out to be indispensible to give a correct treatment of the Delaunay and Voronoi complexes using inverse stereographic projection onto a sphere and to prove rigorously that the well known projective map between the sphere and the paraboloid maps the Delaunay triangulation and the Voronoi diagram w.r.t. the sphere to the more traditional Delaunay triangulation and Voronoi diagram w.r.t. the paraboloid. To the best of our knowledge, this notion of projective polyhedron is new. We also believe that some of our proofs establishing the equivalence of $\mathcal{V}$-polyhedra and $\mathcal{H}$-polyhedra are new.

Chapter 6 on combinatorial topology is hardly original. However, most texts covering this material are either old fashion or too advanced. Yet, this material is used extensively in meshing and geometric modeling. We tried to give a rather intuitive yet rigorous exposition. We decided to introduce the terminology *combinatorial manifold*, a notion usually referred to as *triangulated manifold*.

A recurring theme in these notes is the process of "conification" (algebraically, "homogenization"), that is, forming a cone from some geometric object. Indeed, "conification" turns an object into a set of lines, and since lines play the role of points in projective geometry, "conification" ("homogenization") is the way to "projectivize" geometric affine objects. Then, these (affine) objects appear as "conic sections" of cones by hyperplanes, just the way the classical conics (ellipse, hyperbola, parabola) appear as conic sections.

It is worth warning our readers that convexity and polytope theory is deceptively simple. This is a subject where most intuitive propositions fail as soon as the dimension of the space is greater than 3 (definitely 4), because our human intuition is not very good in dimension greater than 3. Furthermore, rigorous proofs of seemingly very simple facts are often quite complicated and may require sophisticated tools (for example, shellings, for a correct proof of the Euler-Poincaré formula). Nevertheless, readers are urged to strenghten their geometric intuition; they should just be very vigilant! This is another case where Tate's famous saying is more than pertinent: "Reason geometrically, prove algebraically."

At first, these notes were meant as a complement to Chapter 3 (Properties of Convex Sets: A Glimpse) of my book (*Geometric Methods and Applications*, [20]). However, they turn out to cover much more material. For the reader's convenience, I have included Chapter 3 of my book as part of Chapter 2 of these notes. I also assume some familiarity with affine geometry. The reader may wish to review the basics of affine geometry. These can be found in any standard geometry text (Chapter 2 of Gallier [20] covers more than needed for these notes).

Most of the material on convex sets is taken from Berger [6] (*Geometry II*). Other relevant sources include Ziegler [45], Grünbaum [24] Valentine [43], Barvinok [3], Rockafellar [34], Bourbaki (Topological Vector Spaces) [9] and Lax [26], the last four dealing with affine spaces of infinite dimension. As to polytopes and polyhedra, "the" classic reference is Grünbaum [24]. Other good references include Ziegler [45], Ewald [18], Cromwell [14] and Thomas [40].

The recent book by Thomas contains an excellent and easy going presentation of polytope theory. This book also gives an introduction to the theory of triangulations of point configurations, including the definition of secondary polytopes and state polytopes, which happen to play a role in certain areas of biology. For this, a quick but very efficient presentation of Gröbner bases is provided. We highly recommend Thomas's book [40] as further reading. It is also an excellent preparation for the more advanced book by Sturmfels [39]. However, in our opinion, the "bible" on polytope theory is without any contest, Ziegler [45], a masterly and beautiful piece of mathematics. In fact, our Chapter 7 is heavily inspired by Chapter 8 of Ziegler. However, the pace of Ziegler's book is quite brisk and we hope that our more pedestrian account will inspire readers to go back and read the masters.

In a not too distant future, I would like to write about constrained Delaunay triangulations, a formidable topic, please be patient!

I wish to thank Marcelo Siqueira for catching many typos and mistakes and for his many helpful suggestions regarding the presentation. At least a third of this manuscript was written while I was on sabbatical at INRIA, Sophia Antipolis, in the Asclepios Project. My deepest thanks to Nicholas Ayache and his colleagues (especially Xavier Pennec and Hervé Delingette) for inviting me to spend a wonderful and very productive year and for making me feel perfectly at home within the Asclepios Project.

Chapter 2

Basic Properties of Convex Sets

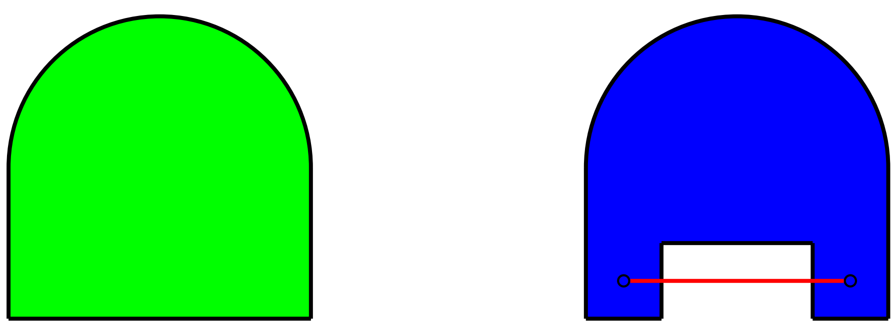
2.1 Convex Sets

Convex sets play a very important role in geometry. In this chapter we state and prove some of the "classics" of convex affine geometry: Carathéodory's theorem, Radon's theorem, and Helly's theorem. These theorems share the property that they are easy to state, but they are deep, and their proof, although rather short, requires a lot of creativity.

Given an affine space E, recall that a subset V of E is *convex* if for any two points $a, b \in V$, we have $c \in V$ for every point $c = (1-\lambda)a + \lambda b$, with $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$). Given any two points a, b, the notation $[a, b]$ is often used to denote the line segment between a and b, that is,

$$[a, b] = \{c \in E \mid c = (1-\lambda)a + \lambda b,\ 0 \leq \lambda \leq 1\},$$

and thus a set V is convex if $[a, b] \subseteq V$ for any two points $a, b \in V$ ($a = b$ is allowed). The empty set is trivially convex, every one-point set $\{a\}$ is convex, and the entire affine space E is of course convex.



(a) (b)

Figure 2.1: (a) A convex set; (b) A nonconvex set

It is obvious that the intersection of any family (finite or infinite) of convex sets is convex. Then, given any (nonempty) subset S of E, there is a smallest convex set containing S denoted by $\mathcal{C}(S)$ or $\mathrm{conv}(S)$ and called the *convex hull of* S (namely, the intersection of all convex sets containing S). The *affine hull* of a subset, S, of E is the smallest affine set containing S and it will be denoted by $\langle S \rangle$ or $\mathrm{aff}(S)$.

Definition 2.1 Given any affine space, E, the *dimension* of a nonempty convex subset, S, of E, denoted by $\dim S$, is the dimension of the smallest affine subset, $\mathrm{aff}(S)$, containing S.

A good understanding of what $\mathcal{C}(S)$ is, and good methods for computing it, are essential. First, we have the following simple but crucial lemma:

Lemma 2.1 *Given an affine space* $\langle E, \overrightarrow{E}, + \rangle$, *for any family* $(a_i)_{i \in I}$ *of points in* E, *the set* V *of convex combinations* $\sum_{i \in I} \lambda_i a_i$ *(where* $\sum_{i \in I} \lambda_i = 1$ *and* $\lambda_i \geq 0$*) is the convex hull of* $(a_i)_{i \in I}$.

Proof. If $(a_i)_{i \in I}$ is empty, then $V = \emptyset$, because of the condition $\sum_{i \in I} \lambda_i = 1$. As in the case of affine combinations, it is easily shown by induction that any convex combination can be obtained by computing convex combinations of two points at a time. As a consequence, if $(a_i)_{i \in I}$ is nonempty, then the smallest convex subspace containing $(a_i)_{i \in I}$ must contain the set V of all convex combinations $\sum_{i \in I} \lambda_i a_i$. Thus, it is enough to show that V is closed under convex combinations, which is immediately verified. $\square$

In view of Lemma 2.1, it is obvious that any affine subspace of E is convex. Convex sets also arise in terms of hyperplanes. Given a hyperplane H, if $f \colon E \to \mathbb{R}$ is any nonconstant affine form defining H (i.e., $H = \mathrm{Ker}\, f$), we can define the two subsets

$$H_+(f) = \{a \in E \mid f(a) \geq 0\} \quad \text{and} \quad H_-(f) = \{a \in E \mid f(a) \leq 0\},$$

called *(closed) half-spaces associated with* f.

Observe that if $\lambda > 0$, then $H_+(\lambda f) = H_+(f)$, but if $\lambda < 0$, then $H_+(\lambda f) = H_-(f)$, and similarly for $H_-(\lambda f)$. However, the set

$$\{H_+(f), H_-(f)\}$$

depends only on the hyperplane H, and the choice of a specific f defining H amounts to the choice of one of the two half-spaces. For this reason, we will also say that $H_+(f)$ and $H_-(f)$ are the *closed half-spaces associated with* H. Clearly, $H_+(f) \cup H_-(f) = E$ and $H_+(f) \cap H_-(f) = H$. It is immediately verified that $H_+(f)$ and $H_-(f)$ are convex. Bounded convex sets arising as the intersection of a finite family of half-spaces associated with hyperplanes play a major role in convex geometry and topology (they are called *convex polytopes*).

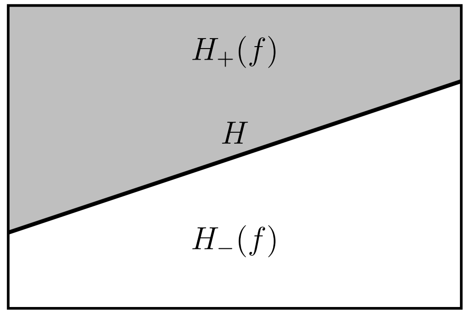


Figure 2.2: The two half-spaces determined by a hyperplane, H

It is natural to wonder whether Lemma 2.1 can be sharpened in two directions: (1) Is it possible to have a fixed bound on the number of points involved in the convex combinations? (2) Is it necessary to consider convex combinations of all points, or is it possible to consider only a subset with special properties?

The answer is yes in both cases. In case 1, assuming that the affine space E has dimension m, Carathéodory's theorem asserts that it is enough to consider convex combinations of $m+1$ points. For example, in the plane $\mathbb{A}^2$, the convex hull of a set S of points is the union of all triangles (interior points included) with vertices in S. In case 2, the theorem of Krein and Milman asserts that a convex set that is also compact is the convex hull of its extremal points (given a convex set S, a point $a \in S$ is extremal if $S - \{a\}$ is also convex, see Berger [6] or Lang [25]). Next, we prove Carathéodory's theorem.

2.2 Carathéodory's Theorem

The proof of Carathéodory's theorem is really beautiful. It proceeds by contradiction and uses a minimality argument.

Theorem 2.2 *(Carathéodory, 1907) Given any affine space E of dimension m, for any (nonvoid) family $S = (a_i)_{i \in L}$ in E, the convex hull $\mathcal{C}(S)$ of S is equal to the set of convex combinations of families of $m+1$ points of S.*

Proof. By Lemma 2.1,

$$\mathcal{C}(S) = \left\{ \sum_{i \in I} \lambda_i a_i \mid a_i \in S, \sum_{i \in I} \lambda_i = 1, \lambda_i \geq 0, I \subseteq L, I \text{ finite} \right\}.$$

We would like to prove that

$$\mathcal{C}(S) = \left\{ \sum_{i \in I} \lambda_i a_i \mid a_i \in S, \sum_{i \in I} \lambda_i = 1, \lambda_i \geq 0, I \subseteq L, |I| = m+1 \right\}.$$

We proceed by contradiction. If the theorem is false, there is some point $b \in \mathcal{C}(S)$ such that b can be expressed as a convex combination $b = \sum_{i\in I} \lambda_i a_i$, where $I \subseteq L$ is a finite set of cardinality $|I| = q$ with $q \geq m+2$, and b cannot be expressed as any convex combination $b = \sum_{j\in J} \mu_j a_j$ of strictly fewer than q points in S, that is, where $|J| < q$. Such a point $b \in \mathcal{C}(S)$ is a convex combination

$$b = \lambda_1 a_1 + \cdots + \lambda_q a_q,$$

where $\lambda_1 + \cdots + \lambda_q = 1$ and $\lambda_i > 0$ $(1 \leq i \leq q)$. We shall prove that b can be written as a convex combination of $q-1$ of the a_i. Pick any origin O in E. Since there are $q > m+1$ points $a_1, \ldots, a_q$, these points are affinely dependent, and by Lemma 2.6.5 from Gallier [20], there is a family $(\mu_1, \ldots, \mu_q)$ all scalars not all null, such that $\mu_1 + \cdots + \mu_q = 0$ and

$$\sum_{i=1}^{q} \mu_i \mathbf{Oa_i} = 0.$$

Consider the set $T \subseteq \mathbb{R}$ defined by

$$T = \{t \in \mathbb{R} \mid \lambda_i + t\mu_i \geq 0,\ \mu_i \neq 0,\ 1 \leq i \leq q\}.$$

The set T is nonempty, since it contains 0. Since $\sum_{i=1}^{q} \mu_i = 0$ and the μ_i are not all null, there are some μ_h, μ_k such that $\mu_h < 0$ and $\mu_k > 0$, which implies that $T = [\alpha, \beta]$, where

$$\alpha = \max_{1\leq i\leq q}\{-\lambda_i/\mu_i \mid \mu_i > 0\} \quad \text{and} \quad \beta = \min_{1\leq i\leq q}\{-\lambda_i/\mu_i \mid \mu_i < 0\}$$

(T is the intersection of the closed half-spaces $\{t \in \mathbb{R} \mid \lambda_i + t\mu_i \geq 0,\ \mu_i \neq 0\}$). Observe that $\alpha < 0 < \beta$, since $\lambda_i > 0$ for all $i = 1, \ldots, q$.

We claim that there is some j $(1 \leq j \leq q)$ such that

$$\lambda_j + \alpha\mu_j = 0.$$

Indeed, since

$$\alpha = \max_{1\leq i\leq q}\{-\lambda_i/\mu_i \mid \mu_i > 0\},$$

as the set on the right hand side is finite, the maximum is achieved and there is some index j so that $\alpha = -\lambda_j/\mu_j$. If j is some index such that $\lambda_j + \alpha\mu_j = 0$, since $\sum_{i=1}^{q} \mu_i \mathbf{Oa_i} = 0$, we have

$$\begin{aligned}
b = \sum_{i=1}^{q} \lambda_i a_i &= O + \sum_{i=1}^{q} \lambda_i \mathbf{Oa_i} + 0, \\
&= O + \sum_{i=1}^{q} \lambda_i \mathbf{Oa_i} + \alpha\left(\sum_{i=1}^{q} \mu_i \mathbf{Oa_i}\right), \\
&= O + \sum_{i=1}^{q} (\lambda_i + \alpha\mu_i)\mathbf{Oa_i}, \\
&= \sum_{i=1}^{q} (\lambda_i + \alpha\mu_i) a_i, \\
&= \sum_{i=1,\, i\neq j}^{q} (\lambda_i + \alpha\mu_i) a_i,
\end{aligned}$$

since $\lambda_j + \alpha\mu_j = 0$. Since $\sum_{i=1}^{q} \mu_i = 0$, $\sum_{i=1}^{q} \lambda_i = 1$, and $\lambda_j + \alpha\mu_j = 0$, we have

$$\sum_{i=1,\, i\neq j}^{q} \lambda_i + \alpha\mu_i = 1,$$

and since $\lambda_i + \alpha\mu_i \geq 0$ for $i = 1, \ldots, q$, the above shows that b can be expressed as a convex combination of $q - 1$ points from S. However, this contradicts the assumption that b cannot be expressed as a convex combination of strictly fewer than q points from S, and the theorem is proved. $\square$

If S is a finite (of infinite) set of points in the affine plane $\mathbb{A}^2$, Theorem 2.2 confirms our intuition that $\mathcal{C}(S)$ is the union of triangles (including interior points) whose vertices belong to S. Similarly, the convex hull of a set S of points in $\mathbb{A}^3$ is the union of tetrahedra (including interior points) whose vertices belong to S. We get the feeling that triangulations play a crucial role, which is of course true!

An interesting consequence of Carathéodory's theorem is the following result:

Proposition 2.3 *If K is any compact subset of $\mathbb{A}^m$, then the convex hull,* $\mathrm{conv}(K)$*, of K is also compact.*

Proposition 2.3 can be proved by showing that $\mathrm{conv}(K)$ is the image of some compact subset of $\mathbb{R}^{m+1} \times (\mathbb{A}^m)^{m+1}$ by some well chosen continuous map.

A closer examination of the proof of Theorem 2.2 reveals that the fact that the μ_i's add up to zero is actually not needed in the proof. This fact ensures that T is a closed interval but all we need is that T be bounded from below, and this only requires that some μ_j be strictly positive. As a consequence, we can prove a version of Theorem 2.2 for convex cones. This is a useful result since cones play such an important role in convex optimization. let us recall some basic definitions about cones.

Definition 2.2 Given any vector space, E, a subset, $C \subseteq E$, is a *convex cone* iff C is closed under *positive linear combinations*, that is, linear combinations of the form,

$$\sum_{i\in I} \lambda_i v_i, \quad \text{with} \quad v_i \in C \quad \text{and} \quad \lambda_i \geq 0 \quad \text{for all} \quad i \in I,$$

where I has finite support (all $\lambda_i = 0$ except for finitely many $i \in I$). Given any set of vectors, S, the *positive hull* of S, or *cone* spanned by S, denoted $\mathrm{cone}(S)$, is the set of all positive linear combinations of vectors in S,

$$\mathrm{cone}(S) = \left\{ \sum_{i\in I} \lambda_i v_i \mid v_i \in S,\ \lambda_i \geq 0 \right\}.$$

Note that a cone always contains 0. When S consists of a finite number of vector, the convex cone, $\mathrm{cone}(S)$, is called a *polyhedral cone*. We have the following version of Carathéodory's theorem for convex cones:

Theorem 2.4 *Given any vector space, E, of dimension m, for any (nonvoid) family $S = (v_i)_{i\in L}$ of vectors in E, the cone, $\mathrm{cone}(S)$, spanned by S is equal to the set of positive combinations of families of m vectors in S.*

The proof of Theorem 2.4 can be easily adapted from the proof of Theorem 2.2 and is left as an exercise.

There is an interesting generalization of Carathéodory's theorem known as the *Colorful Carathéodory theorem*. This theorem due to Bárány and proved in 1982 can be used to give a fairly short proof of a generalization of Helly's theorem known as Tverberg's theorem (see Section 2.4).

Theorem 2.5 *(Colorful Carathéodory theorem) Let E be any affine space of dimension m. For any point, $b \in E$, for any sequence of $m+1$ nonempty subsets, $(S_1, \ldots, S_{m+1})$, of E, if $b \in \mathrm{conv}(S_i)$ for $i = 1, \ldots, m+1$, then there exists a sequence of $m+1$ points, $(a_1, \ldots, a_{m+1})$, with $a_i \in S_i$, so that $b \in \mathrm{conv}(a_1, \ldots, a_{m+1})$, that is, b is a convex combination of the a_i's.*

Although Theorem 2.5 is not hard to prove, we will not prove it here. Instead, we refer the reader to Matousek [27], Chapter 8, Section 8.2. There is also a stronger version of Theorem 2.5, in which it is enough to assume that $b \in \mathrm{conv}(S_i \cup S_j)$ for all i, j with $1 \leq i < j \leq m+1$.

Now that we have given an answer to the first question posed at the end of Section 2.1 we give an answer to the second question.

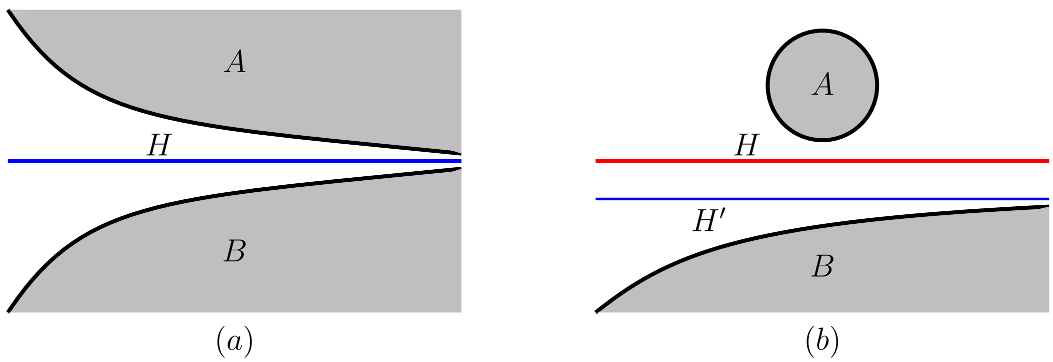


Figure 2.3: (a) A separating hyperplane, H. (b) A strictly separating hyperplane, H

2.3 Vertices, Extremal Points and Krein and Milman's Theorem

First, we define the notions of separation and of separating hyperplanes. For this, recall the definition of the closed (or open) half–spaces determined by a hyperplane.

Given a hyperplane H, if $f\colon E \to \mathbb{R}$ is any nonconstant affine form defining H (i.e., $H = \mathrm{Ker}\, f$), we define the *closed half-spaces associated with* f by

$$
\begin{aligned}
H_+(f) &= \{a \in E \mid f(a) \geq 0\},\\
H_-(f) &= \{a \in E \mid f(a) \leq 0\}.
\end{aligned}
$$

Observe that if $\lambda > 0$, then $H_+(\lambda f) = H_+(f)$, but if $\lambda < 0$, then $H_+(\lambda f) = H_-(f)$, and similarly for $H_-(\lambda f)$.

Thus, the set $\{H_+(f), H_-(f)\}$ depends only on the hyperplane, H, and the choice of a specific f defining H amounts to the choice of one of the two half-spaces.

We also define the *open half–spaces associated with* f as the two sets

$$
\begin{aligned}
\overset{\circ}{H}_+(f) &= \{a \in E \mid f(a) > 0\},\\
\overset{\circ}{H}_-(f) &= \{a \in E \mid f(a) < 0\}.
\end{aligned}
$$

The set $\{\overset{\circ}{H}_+(f), \overset{\circ}{H}_-(f)\}$ only depends on the hyperplane H. Clearly, we have $\overset{\circ}{H}_+(f) = H_+(f) - H$ and $\overset{\circ}{H}_-(f) = H_-(f) - H$.

Definition 2.3 Given an affine space, X, and two nonempty subsets, A and B, of X, we say that a hyperplane H *separates (resp. strictly separates)* A *and* B if A is in one and B is in the other of the two half–spaces (resp. open half–spaces) determined by H.

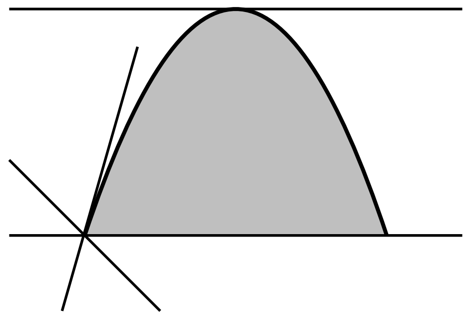


Figure 2.4: Examples of supporting hyperplanes

In Figure 2.3 (a), the two closed convex sets A and B are unbounded and both asymptotic to the hyperplane, H. The hyperplane, H, is a separating hyperplane for A and B but A and B can't be strictly separated. In Figure 2.3 (b), both A and B are convex and closed, B is unbounded and asymptotic to the hyperplane, H', but A is bounded. The hyperplane, H strictly separates A and B. The hyperplane H' also separates A and B but not strictly.

The special case of separation where A is convex and $B = \{a\}$, for some point, a, in A, is of particular importance.

Definition 2.4 Let X be an affine space and let A be any nonempty subset of X. A *supporting hyperplane of* A is any hyperplane, H, containing some point, a, of A, and separating $\{a\}$ and A. We say that H is a *supporting hyperplane of* A *at* a.

Observe that if H is a supporting hyperplane of A at a, then we must have $a \in \partial A$. Otherwise, there would be some open ball $B(a, \epsilon)$ of center a contained in A and so there would be points of A (in $B(a, \epsilon)$) in both half-spaces determined by H, contradicting the fact that H is a supporting hyperplane of A at a. Furthermore, $H \cap \overset{\circ}{A} = \emptyset$.

One should experiment with various pictures and realize that supporting hyperplanes at a point may not exist (for example, if A is not convex), may not be unique, and may have several distinct supporting points! (See Figure 2.4).

Next, we need to define various types of boundary points of closed convex sets.

Definition 2.5 Let X be an affine space of dimension d. For any nonempty closed and convex subset, A, of dimension d, a point $a \in \partial A$ has *order* $k(a)$ if the intersection of all the supporting hyperplanes of A at a is an affine subspace of dimension $k(a)$. We say that $a \in \partial A$ is a *vertex* if $k(a) = 0$; we say that a is *smooth* if $k(a) = d - 1$, i.e., if the supporting hyperplane at a is unique.

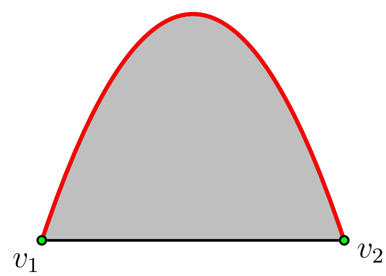


Figure 2.5: Examples of vertices and extreme points

A vertex is a boundary point, a, such that there are d independent supporting hyperplanes at a. A d-simplex has boundary points of order $0, 1, \ldots, d-1$. The following proposition is shown in Berger [6] (Proposition 11.6.2):

Proposition 2.6 *The set of vertices of a closed and convex subset is countable.*

Another important concept is that of an extremal point.

Definition 2.6 Let X be an affine space. For any nonempty convex subset, A, a point $a \in \partial A$ is *extremal* (or *extreme*) if $A - \{a\}$ is still convex.

It is fairly obvious that a point $a \in \partial A$ is extremal if it does not belong to the interior of any closed nontrivial line segment $[x, y] \subseteq A$ ($x \neq y$, $a \neq x$ and $a \neq y$).

Observe that a vertex is extremal, but the converse is false. For example, in Figure 2.5, all the points on the arc of parabola, including v_1 and v_2, are extreme points. However, only v_1 and v_2 are vertices. Also, if $\dim X \geq 3$, the set of extremal points of a compact convex may not be closed.

Actually, it is not at all obvious that a nonempty compact convex set possesses extremal points. In fact, a stronger results holds (Krein and Milman's theorem). In preparation for the proof of this important theorem, observe that any compact (nontrivial) interval of $\mathbb{A}^1$ has two extremal points, its two endpoints. We need the following lemma:

Lemma 2.7 *Let X be an affine space of dimension n, and let A be a nonempty compact and convex set. Then, $A = \mathcal{C}(\partial A)$, i.e., A is equal to the convex hull of its boundary.*

Proof. Pick any a in A, and consider any line, D, through a. Then, $D \cap A$ is closed and convex. However, since A is compact, it follows that $D \cap A$ is a closed interval $[u, v]$ containing a, and $u, v \in \partial A$. Therefore, $a \in \mathcal{C}(\partial A)$, as desired. $\square$

The following important theorem shows that only extremal points matter as far as determining a compact and convex subset from its boundary. The proof of Theorem 2.8 makes use of a proposition due to Minkowski (Proposition 3.18) which will be proved in Section 3.2.

Theorem 2.8 *(Krein and Milman, 1940) Let X be an affine space of dimension n. Every compact and convex nonempty subset, A, is equal to the convex hull of its set of extremal points.*

Proof. Denote the set of extremal points of A by $\mathrm{Extrem}(A)$. We proceed by induction on $d = \dim X$. When $d = 1$, the convex and compact subset A must be a closed interval $[u, v]$, or a single point. In either cases, the theorem holds trivially. Now, assume $d \geq 2$, and assume that the theorem holds for $d - 1$. It is easily verified that

$$\mathrm{Extrem}(A \cap H) = (\mathrm{Extrem}(A)) \cap H,$$

for every supporting hyperplane H of A (such hyperplanes exist, by Minkowski's proposition (Proposition 3.18)). Observe that Lemma 2.7 implies that if we can prove that

$$\partial A \subseteq \mathcal{C}(\mathrm{Extrem}(A)),$$

then, since $A = \mathcal{C}(\partial A)$, we will have established that

$$A = \mathcal{C}(\mathrm{Extrem}(A)).$$

Let $a \in \partial A$, and let H be a supporting hyperplane of A at a (which exists, by Minkowski's proposition). Now, A and H are convex so $A \cap H$ is convex; H is closed and A is compact, so $H \cap A$ is a closed subset of a compact subset, A, and thus, $A \cap H$ is also compact. Since $A \cap H$ is a compact and convex subset of H and H has dimension $d - 1$, by the induction hypothesis, we have

$$A \cap H = \mathcal{C}(\mathrm{Extrem}(A \cap H)).$$

However,

$$\begin{aligned} \mathcal{C}(\mathrm{Extrem}(A \cap H)) &= \mathcal{C}((\mathrm{Extrem}(A)) \cap H) \\ &= \mathcal{C}(\mathrm{Extrem}(A)) \cap H \subseteq \mathcal{C}(\mathrm{Extrem}(A)), \end{aligned}$$

and so, $a \in A \cap H \subseteq \mathcal{C}(\mathrm{Extrem}(A))$. Therefore, we proved that

$$\partial A \subseteq \mathcal{C}(\mathrm{Extrem}(A)),$$

from which we deduce that $A = \mathcal{C}(\mathrm{Extrem}(A))$, as explained earlier. $\square$

Remark: Observe that Krein and Milman's theorem implies that any nonempty compact and convex set has a nonempty subset of extremal points. This is intuitively obvious, but hard to prove! Krein and Milman's theorem also applies to infinite dimensional affine spaces, provided that they are locally convex, see Valentine [43], Chapter 11, Bourbaki [9], Chapter II, Barvinok [3], Chapter 3, or Lax [26], Chapter 13.

An important consequence of Krein and Millman's theorem is that every convex function on a convex and compact set achieves its maximum at some extremal point.

Definition 2.7 Let A be a nonempty convex subset of $\mathbb{A}^n$. A function, $f\colon A \to \mathbb{R}$, is *convex* if

$$f((1-\lambda)a + \lambda b) \leq (1-\lambda)f(a) + \lambda f(b)$$

for all $a, b \in A$ and for all $\lambda \in [0,1]$. The function, $f\colon A \to \mathbb{R}$, is *strictly convex* if

$$f((1-\lambda)a + \lambda b) < (1-\lambda)f(a) + \lambda f(b)$$

for all $a, b \in A$ with $a \neq b$ and for all λ with $0 < \lambda < 1$. A function, $f\colon A \to \mathbb{R}$, is *concave* (resp. *strictly concave*) iff $-f$ is convex (resp. $-f$ is strictly convex).

If f is convex, a simple induction shows that

$$f\left(\sum_{i \in I} \lambda_i a_i\right) \leq \sum_{i \in I} \lambda_i f(a_i)$$

for every finite convex combination in A, *i.e.*, for any finite family $(a_i)_{i \in I}$ of points in A and any family $(\lambda_i)_{i \in I}$ with $\sum_{i \in I} \lambda_i = 1$ and $\lambda_i \geq 0$ for all $i \in I$.

Proposition 2.9 *Let A be a nonempty convex and compact subset of $\mathbb{A}^n$ and let $f\colon A \to \mathbb{R}$ be any function. If f is convex and continuous, then f achieves its maximum at some extreme point of A.*

Proof. Since A is compact and f is continuous, $f(A)$ is a closed interval, $[m, M]$, in $\mathbb{R}$ and so f achieves its minimum m and its maximum M. Say $f(c) = M$, for some $c \in A$. By Krein and Millman's theorem, c is some convex combination of exteme points of A,

$$c = \sum_{i=1}^{k} \lambda_i a_i,$$

with $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \geq 0$ and each a_i an extreme point in A. But then, as f is convex,

$$M = f(c) = f\left(\sum_{i=1}^{k} \lambda_i a_i\right) \leq \sum_{i=1}^{k} \lambda_i f(a_i)$$

and if we let

$$f(a_{i_0}) = \max_{1 \leq i \leq k} \{f(a_i)\}$$

for some i_0 such that $1 \leq i_0 \leq k$, then we get

$$M = f(c) \leq \sum_{i=1}^{k} \lambda_i f(a_i) \leq \left(\sum_{i=1}^{k} \lambda_i\right) f(a_{i_0}) = f(a_{i_0}),$$

as $\sum_{i=1}^{k} \lambda_i = 1$. Since M is the maximum value of the function f over A, we have $f(a_{i_0}) \leq M$ and so,

$$M = f(a_{i_0})$$

and f achieves its maximum at the extreme point, a_{i_0}, as claimed. $\square$

Proposition 2.9 plays an important role in convex optimization: It guarantees that the maximum value of a convex objective function on a compact and convex set is achieved at some extreme point. Thus, it is enough to look for a maximum at some extreme point of the domain.

Proposition 2.9 fails for minimal values of a convex function. For example, the function, $x \mapsto f(x) = x^2$, defined on the compact interval $[-1, 1]$ achieves it minimum at $x = 0$, which is not an extreme point of $[-1, 1]$. However, if f is concave, then f achieves its minimum value at some extreme point of A. In particular, if f is affine, it achieves its minimum and its maximum at some extreme points of A.

We conclude this chapter with three other classics of convex geometry.

2.4 Radon's, Helly's, Tverberg's Theorems and Centerpoints

We begin with *Radon's theorem.*

Theorem 2.10 *(Radon, 1921) Given any affine space E of dimension m, for every subset X of E, if X has at least $m+2$ points, then there is a partition of X into two nonempty disjoint subsets X_1 and X_2 such that the convex hulls of X_1 and X_2 have a nonempty intersection.*

Proof. Pick some origin O in E. Write $X = (x_i)_{i \in L}$ for some index set L (we can let $L = X$). Since by assumption $|X| \geq m+2$ where $m = \dim(E)$, X is affinely dependent, and by Lemma 2.6.5 from Gallier [20], there is a family $(\mu_k)_{k \in L}$ (of finite support) of scalars, not all null, such that

$$\sum_{k \in L} \mu_k = 0 \quad \text{and} \quad \sum_{k \in L} \mu_k \mathbf{O x_k} = 0.$$

Since $\sum_{k \in L} \mu_k = 0$, the μ_k are not all null, and $(\mu_k)_{k \in L}$ has finite support, the sets

$$I = \{i \in L \mid \mu_i > 0\} \quad \text{and} \quad J = \{j \in L \mid \mu_j < 0\}$$

are nonempty, finite, and obviously disjoint. Let

$$X_1 = \{x_i \in X \mid \mu_i > 0\} \quad \text{and} \quad X_2 = \{x_i \in X \mid \mu_i \leq 0\}.$$

Again, since the μ_k are not all null and $\sum_{k \in L} \mu_k = 0$, the sets X_1 and X_2 are nonempty, and obviously

$$X_1 \cap X_2 = \emptyset \quad \text{and} \quad X_1 \cup X_2 = X.$$

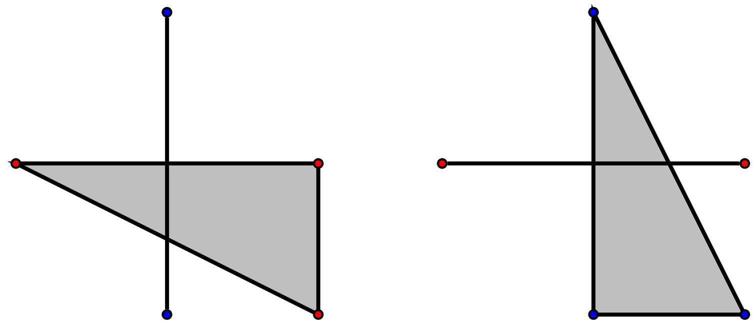


Figure 2.6: Examples of Radon Partitions

Furthermore, the definition of I and J implies that $(x_i)_{i\in I} \subseteq X_1$ and $(x_j)_{j\in J} \subseteq X_2$. It remains to prove that $\mathcal{C}(X_1)\cap\mathcal{C}(X_2) \neq \emptyset$. The definition of I and J implies that

$$\sum_{k\in L} \mu_k \mathbf{Ox_k} = 0$$

can be written as

$$\sum_{i\in I} \mu_i \mathbf{Ox_i} + \sum_{j\in J} \mu_j \mathbf{Ox_j} = 0,$$

that is, as

$$\sum_{i\in I} \mu_i \mathbf{Ox_i} = \sum_{j\in J} -\mu_j \mathbf{Ox_j},$$

where

$$\sum_{i\in I} \mu_i = \sum_{j\in J} -\mu_j = \mu,$$

with $\mu > 0$. Thus, we have

$$\sum_{i\in I} \frac{\mu_i}{\mu} \mathbf{Ox_i} = \sum_{j\in J} -\frac{\mu_j}{\mu} \mathbf{Ox_j},$$

with

$$\sum_{i\in I} \frac{\mu_i}{\mu} = \sum_{j\in J} -\frac{\mu_j}{\mu} = 1,$$

proving that $\sum_{i\in I}(\mu_i/\mu)x_i \in \mathcal{C}(X_1)$ and $\sum_{j\in J} -(\mu_j/\mu)x_j \in \mathcal{C}(X_2)$ are identical, and thus that $\mathcal{C}(X_1)\cap\mathcal{C}(X_2) \neq \emptyset$. $\square$

A partition, (X_1, X_2), of X satisfying the conditions of Theorem 2.10 is sometimes called a *Radon partition* of X and any point in $\mathrm{conv}(X_1)\cap\mathrm{conv}(X_2)$ is called a *Radon point* of X. Figure 2.6 shows two Radon partitions of five points in the plane.

It can be shown that a finite set, $X \subseteq E$, has a unique Radon partition iff it has $m+2$ elements and any $m+1$ points of X are affinely independent. For example, there are exactly two possible cases in the plane as shown in Figure 2.7.

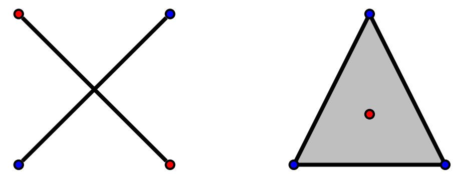


Figure 2.7: The Radon Partitions of four points (in $\mathbb{A}^2$)

There is also a version of Radon's theorem for the class of cones with an apex. Say that a convex cone, $C \subseteq E$, has an *apex* (or is a *pointed cone*) iff there is some hyperplane, H, such that $C \subseteq H_+$ and $H \cap C = \{0\}$. For example, the cone obtained as the intersection of two half spaces in $\mathbb{R}^3$ is not pointed since it is a wedge with a line as part of its boundary. Here is the version of Radon's theorem for convex cones:

Theorem 2.11 *Given any vector space E of dimension m, for every subset X of E, if* $\mathrm{cone}(X)$ *is a pointed cone such that X has at least $m+1$ nonzero vectors, then there is a partition of X into two nonempty disjoint subsets, X_1 and X_2, such that the cones,* $\mathrm{cone}(X_1)$ *and* $\mathrm{cone}(X_2)$*, have a nonempty intersection not reduced to* $\{0\}$.

The proof of Theorem 2.11 is left as an exercise.

There is a beautiful generalization of Radon's theorem known as *Tverberg's Theorem*.

Theorem 2.12 *(Tverberg's Theorem, 1966) Let E be any affine space of dimension m. For any natural number, $r \geq 2$, for every subset, X, of E, if X has at least $(m+1)(r-1)+1$ points, then there is a partition, $(X_1, \ldots, X_r)$, of X into r nonempty pairwise disjoint subsets so that* $\bigcap_{i=1}^{r} \mathrm{conv}(X_i) \neq \emptyset$.

A partition as in Theorem 2.12 is called a *Tverberg partition* and a point in $\bigcap_{i=1}^{r} \mathrm{conv}(X_i)$ is called a *Tverberg point*. Theorem 2.12 was conjectured by Birch and proved by Tverberg in 1966. Tverberg's original proof was technically quite complicated. Tverberg then gave a simpler proof in 1981 and other simpler proofs were later given, notably by Sarkaria (1992) and Onn (1997), using the Colorful Carathéodory theorem. A proof along those lines can be found in Matousek [27], Chapter 8, Section 8.3. A *colored Tverberg theorem* and more can also be found in Matousek [27] (Section 8.3).

Next, we prove a version of *Helly's theorem*.

Theorem 2.13 *(Helly, 1913) Given any affine space E of dimension m, for every family $\{K_1, \ldots, K_n\}$ of n convex subsets of E, if $n \geq m+2$ and the intersection $\bigcap_{i\in I} K_i$ of any $m+1$ of the K_i is nonempty (where $I \subseteq \{1, \ldots, n\}$, $|I| = m+1$), then $\bigcap_{i=1}^{n} K_i$ is nonempty.*

Proof. The proof is by induction on $n \geq m+1$ and uses Radon's theorem in the induction step. For $n = m+1$, the assumption of the theorem is that the intersection of any family of $m+1$ of the K_i's is nonempty, and the theorem holds trivially. Next, let $L = \{1, 2, \ldots, n+1\}$, where $n+1 \geq m+2$. By the induction hypothesis, $C_i = \bigcap_{j \in (L - \{i\})} K_j$ is nonempty for every $i \in L$.

We claim that $C_i \cap C_j \neq \emptyset$ for some $i \neq j$. If so, as $C_i \cap C_j = \bigcap_{k=1}^{n+1} K_k$, we are done. So, let us assume that the C_i's are pairwise disjoint. Then, we can pick a set $X = \{a_1, \ldots, a_{n+1}\}$ such that $a_i \in C_i$, for every $i \in L$. By Radon's Theorem, there are two nonempty disjoint sets $X_1, X_2 \subseteq X$ such that $X = X_1 \cup X_2$ and $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \neq \emptyset$. However, $X_1 \subseteq K_j$ for every j with $a_j \notin X_1$. This is because $a_j \notin K_j$ for every j, and so, we get

$$X_1 \subseteq \bigcap_{a_j \notin X_1} K_j.$$

Symetrically, we also have

$$X_2 \subseteq \bigcap_{a_j \notin X_2} K_j.$$

Since the K_j's are convex and

$$\left(\bigcap_{a_j \notin X_1} K_j\right) \cap \left(\bigcap_{a_j \notin X_2} K_j\right) = \bigcap_{i=1}^{n+1} K_i,$$

it follows that $\mathcal{C}(X_1) \cap \mathcal{C}(X_2) \subseteq \bigcap_{i=1}^{n+1} K_i$, so that $\bigcap_{i=1}^{n+1} K_i$ is nonempty, contradicting the fact that $C_i \cap C_j = \emptyset$ for all $i \neq j$. $\square$

A more general version of Helly's theorem is proved in Berger [6]. An amusing corollary of Helly's theorem is the following result: Consider $n \geq 4$ parallel line segments in the affine plane $\mathbb{A}^2$. If every three of these line segments meet a line, then all of these line segments meet a common line.

We conclude this chapter with a nice application of Helly's Theorem to the existence of centerpoints. Centerpoints generalize the notion of median to higher dimensions. Recall that if we have a set of n data points, $S = \{a_1, \ldots, a_n\}$, on the real line, a *median* for S is a point, x, such that both intervals $[x, \infty)$ and $(-\infty, x]$ contain at least $n/2$ of the points in S (by $n/2$, we mean the largest integer greater than or equal to $n/2$).

Given any hyperplane, H, recall that the closed half-spaces determined by H are denoted H_+ and H_- and that $H \subseteq H_+$ and $H \subseteq H_-$. We let $\overset{\circ}{H}_+ = H_+ - H$ and $\overset{\circ}{H}_- = H_- - H$ be the *open half-spaces* determined by H.

Definition 2.8 Let $S = \{a_1, \ldots, a_n\}$ be a set of n points in $\mathbb{A}^d$. A point, $c \in \mathbb{A}^d$, is a *centerpoint of* S iff for every hyperplane, H, whenever the closed half-space H_+ (resp. H_-) contains c, then H_+ (resp. H_-) contains at least $\frac{n}{d+1}$ points from S (by $\frac{n}{d+1}$, we mean the largest integer greater than or equal to $\frac{n}{d+1}$, namely the ceiling $\lceil \frac{n}{d+1} \rceil$ of $\frac{n}{d+1}$).

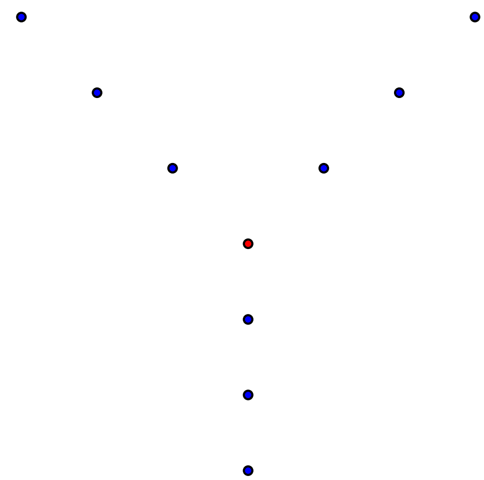


Figure 2.8: Example of a centerpoint

So, for $d = 2$, for each line, D, if the closed half-plane D_+ (resp. D_-) contains c, then D_+ (resp. D_-) contains at least a third of the points from S. For $d = 3$, for each plane, H, if the closed half-space H_+ (resp. H_-) contains c, then H_+ (resp. H_-) contains at least a fourth of the points from S, *etc.* Example 2.8 shows nine points in the plane and one of their centerpoints (in red). This example shows that the bound $\frac{1}{3}$ is tight.

Observe that a point, $c \in \mathbb{A}^d$, is a centerpoint of S iff c belongs to every open half-space, $\overset{\circ}{H}_+$ (resp. $\overset{\circ}{H}_-$) containing at least $\frac{dn}{d+1} + 1$ points from S (again, we mean $\lceil \frac{dn}{d+1} \rceil + 1$).

Indeed, if c is a centerpoint of S and H is any hyperplane such that $\overset{\circ}{H}_+$ (resp. $\overset{\circ}{H}_-$) contains at least $\frac{dn}{d+1} + 1$ points from S, then $\overset{\circ}{H}_+$ (resp. $\overset{\circ}{H}_-$) must contain c as otherwise, the closed half-space, H_- (resp. H_+) would contain c and at most $n - \frac{dn}{d+1} - 1 = \frac{n}{d+1} - 1$ points from S, a contradiction. Conversely, assume that c belongs to every open half-space, $\overset{\circ}{H}_+$ (resp. $\overset{\circ}{H}_-$) containing at least $\frac{dn}{d+1} + 1$ points from S. Then, for any hyperplane, H, if $c \in H_+$ (resp. $c \in H_-$) but H_+ contains at most $\frac{n}{d+1} - 1$ points from S, then the open half-space, $\overset{\circ}{H}_-$ (resp. $\overset{\circ}{H}_+$) would contain at least $n - \frac{n}{d+1} + 1 = \frac{dn}{d+1} + 1$ points from S but not c, a contradiction.

We are now ready to prove the existence of centerpoints.

Theorem 2.14 *(Existence of Centerpoints) Every finite set, $S = \{a_1, \ldots, a_n\}$, of n points in $\mathbb{A}^d$ has some centerpoint.*

Proof. We will use the second characterization of centerpoints involving open half-spaces containing at least $\frac{dn}{d+1} + 1$ points.

Consider the family of sets,

$$\begin{aligned}\mathcal{C} \;=\; & \left\{\mathrm{conv}(S \cap \overset{\circ}{H}_+) \mid (\exists H) \left(|S \cap \overset{\circ}{H}_+ | > \frac{dn}{d+1}\right)\right\} \\ & \cup \left\{\mathrm{conv}(S \cap \overset{\circ}{H}_-) \mid (\exists H) \left(|S \cap \overset{\circ}{H}_- | > \frac{dn}{d+1}\right)\right\},\end{aligned}$$

where H is a hyperplane.

As S is finite, $\mathcal{C}$ consists of a finite number of convex sets, say $\{C_1, \ldots, C_m\}$. If we prove that $\bigcap_{i=1}^m C_i \neq \emptyset$ we are done, because $\bigcap_{i=1}^m C_i$ is the set of centerpoints of S.

First, we prove by induction on k (with $1 \leq k \leq d+1$), that any intersection of k of the C_i's has at least $\frac{(d+1-k)n}{d+1} + k$ elements from S. For $k = 1$, this holds by definition of the C_i's.

Next, consider the intersection of $k+1 \leq d+1$ of the C_i's, say $C_{i_1} \cap \cdots \cap C_{i_k} \cap C_{i_{k+1}}$. Let

$$\begin{aligned} A &= S \cap (C_{i_1} \cap \cdots \cap C_{i_k} \cap C_{i_{k+1}}) \\ B &= S \cap (C_{i_1} \cap \cdots \cap C_{i_k}) \\ C &= S \cap C_{i_{k+1}}. \end{aligned}$$

Note that $A = B \cap C$. By the induction hypothesis, B contains at least $\frac{(d+1-k)n}{d+1} + k$ elements from S. As C contains at least $\frac{dn}{d+1} + 1$ points from S, and as

$$|B \cup C| = |B| + |C| - |B \cap C| = |B| + |C| - |A|$$

and $|B \cup C| \leq n$, we get $n \geq |B| + |C| - |A|$, that is,

$$|A| \geq |B| + |C| - n.$$

It follows that

$$|A| \geq \frac{(d+1-k)n}{d+1} + k + \frac{dn}{d+1} + 1 - n$$

that is,

$$|A| \geq \frac{(d+1-k)n + dn - (d+1)n}{d+1} + k + 1 = \frac{(d+1-(k+1))n}{d+1} + k + 1,$$

establishing the induction hypothesis.

Now, if $m \leq d+1$, the above claim for $k = m$ shows that $\bigcap_{i=1}^m C_i \neq \emptyset$ and we are done. If $m \geq d+2$, the above claim for $k = d+1$ shows that any intersection of $d+1$ of the C_i's is nonempty. Consequently, the conditions for applying Helly's Theorem are satisfied and therefore,

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

However, $\bigcap_{i=1}^{m} C_i$ is the set of centerpoints of S and we are done. $\square$

Remark: The above proof actually shows that the set of centerpoints of S is a convex set. In fact, it is a finite intersection of convex hulls of finitely many points, so it is the convex hull of finitely many points, in other words, a polytope. It should also be noted that Theorem 2.14 can be proved easily using Tverberg's theorem (Theorem 2.12). Indeed, for a judicious choice of r, any Tverberg point is a centerpoint!

Jadhav and Mukhopadhyay have given a linear-time algorithm for computing a centerpoint of a finite set of points in the plane. For $d \geq 3$, it appears that the best that can be done (using linear programming) is $O(n^d)$. However, there are good approximation algorithms (Clarkson, Eppstein, Miller, Sturtivant and Teng) and in $\mathbb{E}^3$ there is a near quadratic algorithm (Agarwal, Sharir and Welzl). Recently, Miller and Sheehy (2009) have given an algorithm for finding an approximate centerpoint in sub-exponential time together with a polynomial-checkable proof of the approximation guarantee.

Chapter 3

Separation and Supporting Hyperplanes

3.1 Separation Theorems and Farkas Lemma

It seems intuitively rather obvious that if A and B are two nonempty disjoint convex sets in $\mathbb{A}^2$, then there is a line, H, separating them, in the sense that A and B belong to the two (disjoint) open half–planes determined by H. However, this is not always true! For example, this fails if both A and B are closed and unbounded (find an example). Nevertheless, the result is true if both A and B are open, or if the notion of separation is weakened a little bit. The key result, from which most separation results follow, is a geometric version of the *Hahn-Banach theorem*. In the sequel, we restrict our attention to real affine spaces of finite dimension. Then, if X is an affine space of dimension d, there is an affine bijection f between X and $\mathbb{A}^d$.

Now, $\mathbb{A}^d$ is a topological space, under the usual topology on $\mathbb{R}^d$ (in fact, $\mathbb{A}^d$ is a metric space). Recall that if $a = (a_1, \ldots, a_d)$ and $b = (b_1, \ldots, b_d)$ are any two points in $\mathbb{A}^d$, their **Euclidean distance**, $d(a, b)$, is given by

$$d(a, b) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_d - a_d)^2},$$

which is also the *norm*, $\|\mathbf{ab}\|$, of the vector $\mathbf{ab}$ and that for any $\epsilon > 0$, the *open ball of center a and radius ϵ*, $B(a, \epsilon)$, is given by

$$B(a, \epsilon) = \{b \in \mathbb{A}^d \mid d(a, b) < \epsilon\}.$$

A subset $U \subseteq \mathbb{A}^d$ is *open* (in the *norm topology*) if either U is empty or for every point, $a \in U$, there is some (small) open ball, $B(a, \epsilon)$, contained in U. A subset $C \subseteq \mathbb{A}^d$ is *closed* iff $\mathbb{A}^d - C$ is open. For example, the *closed balls*, $\overline{B(a, \epsilon)}$, where

$$\overline{B(a, \epsilon)} = \{b \in \mathbb{A}^d \mid d(a, b) \leq \epsilon\},$$

are closed. A subset $W \subseteq \mathbb{A}^d$ is *bounded* iff there is some ball (open or closed), B, so that $W \subseteq B$. A subset $W \subseteq \mathbb{A}^d$ is *compact* iff every family, $\{U_i\}_{i \in I}$, that is an open cover of W (which means that $W = \bigcup_{i \in I}(W \cap U_i)$, with each U_i an open set) possesses a finite subcover (which means that there is a finite subset, $F \subseteq I$, so that $W = \bigcup_{i \in F}(W \cap U_i)$). In $\mathbb{A}^d$, it can be shown that a subset W is compact iff W is closed and bounded. Given a function, $f\colon \mathbb{A}^m \to \mathbb{A}^n$, we say that f is *continuous* if $f^{-1}(V)$ is open in $\mathbb{A}^m$ whenever V is open in $\mathbb{A}^n$. If $f\colon \mathbb{A}^m \to \mathbb{A}^n$ is a continuous function, although it is generally **false** that $f(U)$ is open if $U \subseteq \mathbb{A}^m$ is open, it is easily checked that $f(K)$ is compact if $K \subseteq \mathbb{A}^m$ is compact.

An affine space X of dimension d becomes a topological space if we give it the topology for which the open subsets are of the form $f^{-1}(U)$, where U is any open subset of $\mathbb{A}^d$ and $f\colon X \to \mathbb{A}^d$ is an affine bijection.

Given any subset, A, of a topological space, X, the smallest closed set containing A is denoted by $\overline{A}$, and is called the *closure* or *adherence of* A. A subset, A, of X, is *dense in* X if $\overline{A} = X$. The largest open set contained in A is denoted by $\overset{\circ}{A}$, and is called the *interior of* A. The set, $\mathrm{Fr}\, A = \overline{A} \cap \overline{X - A}$, is called the *boundary* (or *frontier*) of A. We also denote the boundary of A by ∂A.

In order to prove the Hahn-Banach theorem, we will need two lemmas. Given any two distinct points $x, y \in X$, we let

$$]x, y[= \{(1-\lambda)x + \lambda y \in X \mid 0 < \lambda < 1\}.$$

Our first lemma (Lemma 3.1) is intuitively quite obvious so the reader might be puzzled by the length of its proof. However, after proposing several wrong proofs, we realized that its proof is more subtle than it might appear. The proof below is due to Valentine [43]. See if you can find a shorter (and correct) proof!

Lemma 3.1 *Let S be a nonempty convex set and let $x \in \overset{\circ}{S}$ and $y \in \overline{S}$. Then, we have $]x, y[\subseteq \overset{\circ}{S}$.*

Proof. Let $z \in]x, y[$, that is, $z = (1-\lambda)x + \lambda y$, with $0 < \lambda < 1$. Since $x \in \overset{\circ}{S}$, we can find some open subset, U, contained in S so that $x \in U$. It is easy to check that the central magnification of center z, $H_{z, \frac{\lambda-1}{\lambda}}$, maps x to y. Then, $V = H_{z, \frac{\lambda-1}{\lambda}}(U)$ is an open subset containing y and as $y \in \overline{S}$, we have $V \cap S \neq \emptyset$. Let $v \in V \cap S$ be a point of S in this intersection. Now, there is a unique point, $u \in U \subseteq S$, such that $H_{z, \frac{\lambda-1}{\lambda}}(u) = v$ and, as S is convex, we deduce that $z = (1-\lambda)u + \lambda v \in S$. Since U is open, the set

$$W = (1-\lambda)U + \lambda v = \{(1-\lambda)w + \lambda v \mid w \in U\} \subseteq S$$

is also open and $z \in W$, which shows that $z \in \overset{\circ}{S}$. $\square$

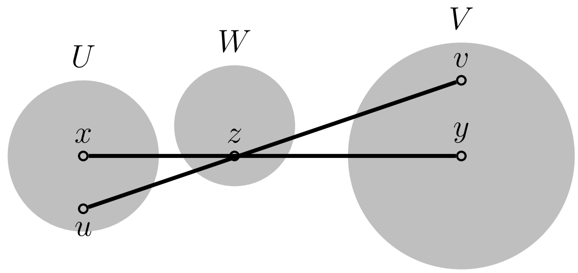


Figure 3.1: Illustration for the proof of Lemma 3.1

Corollary 3.2 *If S is convex, then $\mathring{S}$ is also convex, and we have $\mathring{S} = \mathring{\overline{S}}$. Furthermore, if $\mathring{S} \neq \emptyset$, then $\overline{S} = \overline{\mathring{S}}$.*

Beware that if S is a closed set, then the convex hull, $\mathrm{conv}(S)$, of S is not necessarily closed! (Find a counter-example.) However, if S is compact, then $\mathrm{conv}(S)$ is also compact and thus, closed (see Proposition 2.3).

There is a simple criterion to test whether a convex set has an empty interior, based on the notion of dimension of a convex set (recall that the dimension of a nonempty convex subset is the dimension of its affine hull).

Proposition 3.3 *A nonempty convex set S has a nonempty interior iff* $\dim S = \dim X$.

Proof. Let $d = \dim X$. First, assume that $\mathring{S} \neq \emptyset$. Then, S contains some open ball of center a_0, and in it, we can find a frame $(a_0, a_1, \ldots, a_d)$ for X. Thus, $\dim S = \dim X$. Conversely, let $(a_0, a_1, \ldots, a_d)$ be a frame of X, with $a_i \in S$, for $i = 0, \ldots, d$. Then, we have

$$\frac{a_0 + \cdots + a_d}{d+1} \in \mathring{S},$$

and $\mathring{S}$ is nonempty. $\square$

Proposition 3.3 is false in infinite dimension.

We leave the following property as an exercise:

Proposition 3.4 *If S is convex, then $\overline{S}$ is also convex.*

One can also easily prove that convexity is preserved under direct image and inverse image by an affine map.

The next lemma, which seems intuitively obvious, is the core of the proof of the Hahn-Banach theorem. This is the case where the affine space has dimension two. First, we need to define what is a convex cone with vertex x.

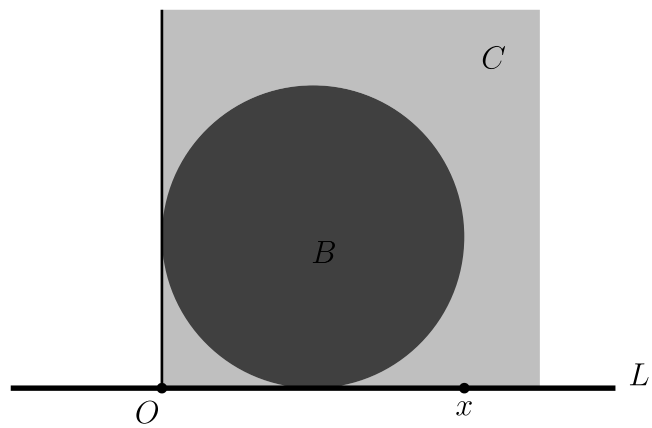


Figure 3.2: Hahn-Banach Theorem in the plane (Lemma 3.5)

Definition 3.1 A convex set, C, is a *convex cone with vertex* x if C is invariant under all central magnifications, $H_{x,\lambda}$, of center x and ratio λ, with $\lambda > 0$ (i.e., $H_{x,\lambda}(C) = C$).

Given a convex set, S, and a point, $x \notin S$, we can define

$$\mathrm{cone}_x(S) = \bigcup_{\lambda > 0} H_{x,\lambda}(S).$$

It is easy to check that this is a convex cone with vertex x.

Lemma 3.5 *Let B be a nonempty open and convex subset of $\mathbb{A}^2$, and let O be a point of $\mathbb{A}^2$ so that $O \notin B$. Then, there is some line, L, through O, so that $L \cap B = \emptyset$.*

Proof. Define the convex cone $C = \mathrm{cone}_O(B)$. As B is open, it is easy to check that each $H_{O,\lambda}(B)$ is open and since C is the union of the $H_{O,\lambda}(B)$ (for $\lambda > 0$), which are open, C itself is open. Also, $O \notin C$. We claim that at least one point, x, of the boundary, ∂C, of C, is distinct from O. Otherwise, $\partial C = \{O\}$ and we claim that $C = \mathbb{A}^2 - \{O\}$, which is not convex, a contradiction. Indeed, as C is convex it is connected, $\mathbb{A}^2 - \{O\}$ itself is connected and $C \subseteq \mathbb{A}^2 - \{O\}$. If $C \neq \mathbb{A}^2 - \{O\}$, pick some point $a \neq O$ in $\mathbb{A}^2 - C$ and some point $c \in C$. Now, a basic property of connectivity asserts that every continuous path from a (in the exterior of C) to c (in the interior of C) must intersect the boundary of C, namely, $\{O\}$. However, there are plenty of paths from a to c that avoid O, a contradiction. Therefore, $C = \mathbb{A}^2 - \{O\}$.

Since C is open and $x \in \partial C$, we have $x \notin C$. Furthermore, we claim that $y = 2O - x$ (the symmetric of x w.r.t. O) does not belong to C either. Otherwise, we would have $y \in \mathring{C} = C$ and $x \in \overline{C}$, and by Lemma 3.1, we would get $O \in C$, a contradiction. Therefore, the line through O and x misses C entirely (since C is a cone), and thus, $B \subseteq C$. $\square$

Finally, we come to the Hahn-Banach theorem.

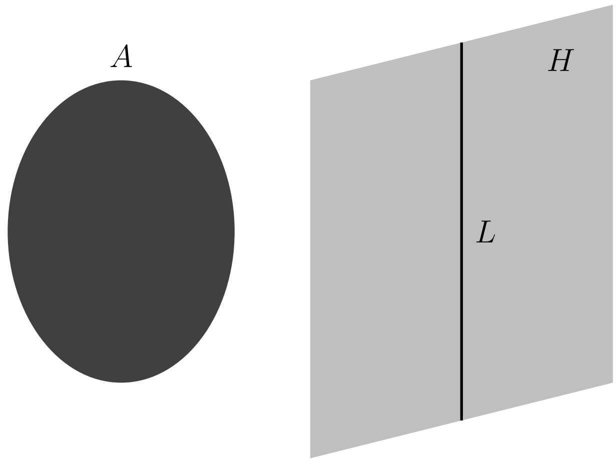


Figure 3.3: Hahn-Banach Theorem, geometric form (Theorem 3.6)

Theorem 3.6 *(Hahn-Banach Theorem, geometric form) Let X be a (finite-dimensional) affine space, A be a nonempty open and convex subset of X and L be an affine subspace of X so that $A \cap L = \emptyset$. Then, there is some hyperplane, H, containing L, that is disjoint from A.*

Proof. The case where $\dim X = 1$ is trivial. Thus, we may assume that $\dim X \geq 2$. We reduce the proof to the case where $\dim X = 2$. Let V be an affine subspace of X of maximal dimension containing L and so that $V \cap A = \emptyset$. Pick an origin $O \in L$ in X, and consider the vector space X_O. We would like to prove that V is a hyperplane, i.e., $\dim V = \dim X - 1$. We proceed by contradiction. Thus, assume that $\dim V \leq \dim X - 2$. In this case, the quotient space X/V has dimension at least 2. We also know that X/V is isomorphic to the orthogonal complement, $V^{\perp}$, of V so we may identify X/V and $V^{\perp}$. The (orthogonal) projection map, $\pi \colon X \to V^{\perp}$, is linear, continuous, and we can show that π maps the open subset A to an open subset $\pi(A)$, which is also convex (one way to prove that $\pi(A)$ is open is to observe that for any point, $a \in A$, a small open ball of center a contained in A is projected by π to an open ball contained in $\pi(A)$ and as π is surjective, $\pi(A)$ is open). Furthermore, $0 \notin \pi(A)$. Since $V^{\perp}$ has dimension at least 2, there is some plane P (a subspace of dimension 2) intersecting $\pi(A)$, and thus, we obtain a nonempty open and convex subset $B = \pi(A) \cap P$ in the plane $P \cong \mathbb{A}^2$. So, we can apply Lemma 3.5 to B and the point $O = 0$ in $P \cong \mathbb{A}^2$ to find a line, l, (in P) through O with $l \cap B = \emptyset$. But then, $l \cap \pi(A) = \emptyset$ and $W = \pi^{-1}(l)$ is an affine subspace such that $W \cap A = \emptyset$ and W properly contains V, contradicting the maximality of V. $\square$

Remark: The geometric form of the Hahn-Banach theorem also holds when the dimension of X is infinite but a slightly more sophisticated proof is required. Actually, all that is needed is to prove that a maximal affine subspace containing L and disjoint from A exists. This can

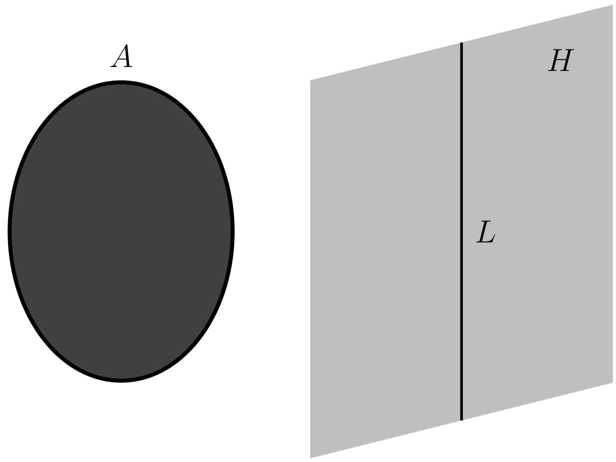


Figure 3.4: Hahn-Banach Theorem, second version (Theorem 3.7)

be done using Zorn's lemma. For other proofs, see Bourbaki [9], Chapter 2, Valentine [43], Chapter 2, Barvinok [3], Chapter 2, or Lax [26], Chapter 3.

Theorem 3.6 is false if we omit the assumption that A is open. For a counter-example, let $A \subseteq \mathbb{A}^2$ be the union of the half space $y < 0$ with the closed segment $[0, 1]$ on the x-axis and let L be the point $(2, 0)$ on the boundary of A. It is also false if A is closed! (Find a counter-example).

Theorem 3.6 has many important corollaries. For example, we will eventually prove that for any two nonempty disjoint convex sets, A and B, there is a hyperplane separating A and B, but this will take some work (recall the definition of a separating hyperplane given in Definition 2.3). We begin with the following version of the Hahn-Banach theorem:

Theorem 3.7 *(Hahn-Banach, second version) Let X be a (finite-dimensional) affine space, A be a nonempty convex subset of X with nonempty interior and L be an affine subspace of X so that $A \cap L = \emptyset$. Then, there is some hyperplane, H, containing L and separating L and A.*

Proof. Since A is convex, by Corollary 3.2, $\mathring{A}$ is also convex. By hypothesis, $\mathring{A}$ is nonempty. So, we can apply Theorem 3.6 to the nonempty open and convex $\mathring{A}$ and to the affine subspace L. We get a hyperplane H containing L such that $\mathring{A} \cap H = \emptyset$. However, $A \subseteq \overline{A} = \overline{\mathring{A}}$ and $\overline{\mathring{A}}$ is contained in the closed half space (H_+ or H_-) containing $\mathring{A}$, so H separates A and L. $\square$

Corollary 3.8 *Given an affine space, X, let A and B be two nonempty disjoint convex subsets and assume that A has nonempty interior ($\mathring{A} \neq \emptyset$). Then, there is a hyperplane separating A and B.*

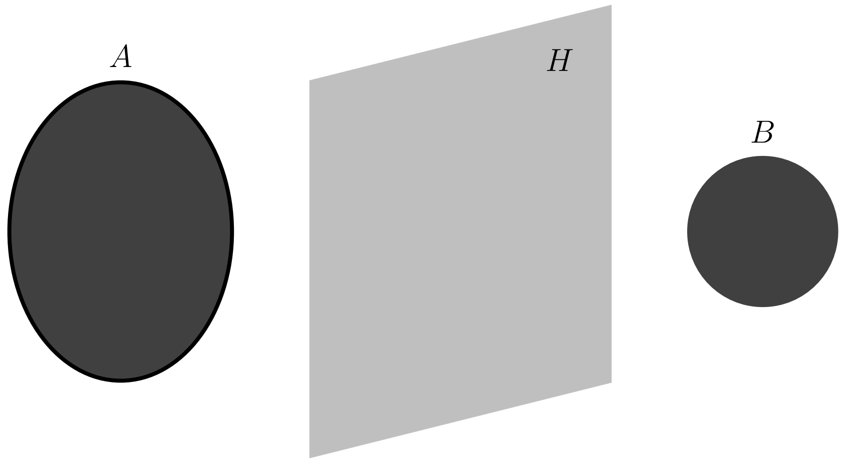


Figure 3.5: Separation Theorem, version 1 (Corollary 3.8)

Proof. Pick some origin O and consider the vector space X_O. Define $C = A - B$ (a special case of the Minkowski sum) as follows:

$$A - B = \{a - b \mid a \in A,\ b \in B\} = \bigcup_{b \in B} (A - b).$$

It is easily verified that $C = A - B$ is convex and has nonempty interior (as a union of subsets having a nonempty interior). Furthermore $O \notin C$, since $A \cap B = \emptyset$.[1] (Note that the definition depends on the choice of O, but this has no effect on the proof.) Since $\overset{\circ}{C}$ is nonempty, we can apply Theorem 3.7 to C and to the affine subspace $\{O\}$ and we get a hyperplane, H, separating C and $\{O\}$. Let f be any linear form defining the hyperplane H. We may assume that $f(a - b) \leq 0$, for all $a \in A$ and all $b \in B$, *i.e.*, $f(a) \leq f(b)$. Consequently, if we let $\alpha = \sup\{f(a) \mid a \in A\}$ (which makes sense, since the set $\{f(a) \mid a \in A\}$ is bounded), we have $f(a) \leq \alpha$ for all $a \in A$ and $f(b) \geq \alpha$ for all $b \in B$, which shows that the affine hyperplane defined by $f - \alpha$ separates A and B. $\square$

Remark: Theorem 3.7 and Corollary 3.8 also hold in the infinite dimensional case, see Lax [26], Chapter 3, or Barvinok, Chapter 3.

Since a hyperplane, H, separating A and B as in Corollary 3.8 is the boundary of each of the two half–spaces that it determines, we also obtain the following corollary:

[1] Readers who prefer a purely affine argument may define $C = A - B$ as the *affine* subset

$$A - B = \{O + a - b \mid a \in A,\ b \in B\}.$$

Again, $O \notin C$ and C is convex. By adjusting O we can pick the affine form, f, defining a separating hyperplane, H, of C and $\{O\}$, so that $f(O + a - b) \leq f(O)$, for all $a \in A$ and all $b \in B$, *i.e.*, $f(a) \leq f(b)$.

Corollary 3.9 *Given an affine space, X, let A and B be two nonempty disjoint open and convex subsets. Then, there is a hyperplane strictly separating A and B.*

Beware that Corollary 3.9 *fails* for *closed* convex sets. However, Corollary 3.9 holds if we also assume that A (or B) is compact.

We need to review the notion of distance from a point to a subset. Let X be a metric space with distance function, d. Given any point, $a \in X$, and any nonempty subset, B, of X, we let

$$d(a, B) = \inf_{b \in B} d(a, b)$$

(where inf is the notation for least upper bound).

Now, if X is an affine space of dimension d, it can be given a metric structure by giving the corresponding vector space a metric structure, for instance, the metric induced by a Euclidean structure. We have the following important property: For any nonempty closed subset, $S \subseteq X$ (not necessarily convex), and any point, $a \in X$, there is some point $s \in S$ "achieving the distance from a to S," i.e., so that

$$d(a, S) = d(a, s).$$

The proof uses the fact that the distance function is continuous and that a continuous function attains its minimum on a compact set, and is left as an exercise.

Corollary 3.10 *Given an affine space, X, let A and B be two nonempty disjoint closed and convex subsets, with A compact. Then, there is a hyperplane strictly separating A and B.*

Proof sketch. First, we pick an origin O and we give $X_O \cong \mathbb{A}^n$ a Euclidean structure. Let d denote the associated distance. Given any subsets A of X, let

$$A + B(O, \epsilon) = \{x \in X \mid d(x, A) < \epsilon\},$$

where $B(a, \epsilon)$ denotes the open ball, $B(a, \epsilon) = \{x \in X \mid d(a, x) < \epsilon\}$, of center a and radius $\epsilon > 0$. Note that

$$A + B(O, \epsilon) = \bigcup_{a \in A} B(a, \epsilon),$$

which shows that $A + B(O, \epsilon)$ is open; furthermore it is easy to see that if A is convex, then $A + B(O, \epsilon)$ is also convex. Now, the function $a \mapsto d(a, B)$ (where $a \in A$) is continuous and since A is compact, it achieves its minimum, $d(A, B) = \min_{a \in A} d(a, B)$, at some point, a, of A. Say, $d(A, B) = \delta$. Since B is closed, there is some $b \in B$ so that $d(A, B) = d(a, B) = d(a, b)$ and since $A \cap B = \emptyset$, we must have $\delta > 0$. Thus, if we pick $\epsilon < \delta/2$, we see that

$$(A + B(O, \epsilon)) \cap (B + B(O, \epsilon)) = \emptyset.$$

Now, $A+B(O,\epsilon)$ and $B+B(O,\epsilon)$ are open, convex and disjoint and we conclude by applying Corollary 3.9. $\square$

A "cute" application of Corollary 3.10 is one of the many versions of "Farkas Lemma" (1893-1894, 1902), a basic result in the theory of linear programming. For any vector, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, and any real, $\alpha \in \mathbb{R}$, write $x \geq \alpha$ iff $x_i \geq \alpha$, for $i = 1, \ldots, n$.

Lemma 3.11 *(Farkas Lemma, Version I) Given any $d \times n$ real matrix, A, and any vector, $z \in \mathbb{R}^d$, exactly one of the following alternatives occurs:*

(a) The linear system, $Ax = z$, has a solution, $x = (x_1, \ldots, x_n)$, such that $x \geq 0$ and $x_1 + \cdots + x_n = 1$, or

(b) There is some $c \in \mathbb{R}^d$ and some $\alpha \in \mathbb{R}$ such that $c^\top z < \alpha$ and $c^\top A \geq \alpha$.

Proof. Let $A_1, \ldots, A_n \in \mathbb{R}^d$ be the n points corresponding to the columns of A. Then, either $z \in \mathrm{conv}(\{A_1, \ldots, A_n\})$ or $z \notin \mathrm{conv}(\{A_1, \ldots, A_n\})$. In the first case, we have a convex combination

$$z = x_1 A_1 + \cdots + x_n A_n$$

where $x_i \geq 0$ and $x_1 + \cdots + x_n = 1$, so $x = (x_1, \ldots, x_n)$ is a solution satisfying (a).

In the second case, by Corollary 3.10, there is a hyperplane, H, strictly separating $\{z\}$ and $\mathrm{conv}(\{A_1, \ldots, A_n\})$, which is obviously closed. In fact, observe that $z \notin \mathrm{conv}(\{A_1, \ldots, A_n\})$ iff there is a hyperplane, H, such that $z \in \overset{\circ}{H}_-$ and $A_i \in H_+$, or $z \in \overset{\circ}{H}_+$ and $A_i \in H_-$, for $i = 1, \ldots, n$. As the affine hyperplane, H, is the zero locus of an equation of the form

$$c_1 y_1 + \cdots + c_d y_d = \alpha,$$

either $c^\top z < \alpha$ and $c^\top A_i \geq \alpha$ for $i = 1, \ldots, n$, that is, $c^\top A \geq \alpha$, or $c^\top z > \alpha$ and $c^\top A \leq \alpha$. In the second case, $(-c)^\top z < -\alpha$ and $(-c)^\top A \geq -\alpha$, so (b) is satisfied by either c and α or by $-c$ and $-\alpha$. $\square$

Remark: If we relax the requirements on solutions of $Ax = z$ and only require $x \geq 0$ ($x_1 + \cdots + x_n = 1$ is no longer required) then, in condition (b), we can take $\alpha = 0$. This is another version of Farkas Lemma. In this case, instead of considering the convex hull of $\{A_1, \ldots, A_n\}$ we are considering the convex cone,

$$\mathrm{cone}(A_1, \ldots, A_n) = \{\lambda A_1 + \cdots + \lambda_n A_n \mid \lambda_i \geq 0,\ 1 \leq i \leq n\},$$

that is, we are dropping the condition $\lambda_1 + \cdots + \lambda_n = 1$. For this version of Farkas Lemma we need the following separation lemma:

Proposition 3.12 *Let $C \subseteq \mathbb{E}^d$ be any closed convex cone with vertex O. Then, for every point, a, not in C, there is a hyperplane, H, passing through O separating a and C with $a \notin H$.*

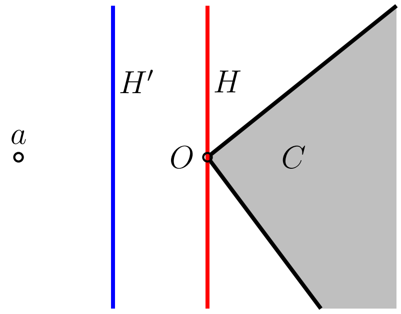


Figure 3.6: Illustration for the proof of Proposition 3.12

Proof. Since C is closed and convex and $\{a\}$ is compact and convex, by Corollary 3.10, there is a hyperplane, H', strictly separating a and C. Let H be the hyperplane through O parallel to H'. Since C and a lie in the two disjoint open half-spaces determined by H', the point a cannot belong to H. Suppose that some point, $b \in C$, lies in the open half-space determined by H and a. Then, the line, L, through O and b intersects H' in some point, c, and as C is a cone, the half line determined by O and b is contained in C. So, $c \in C$ would belong to H', a contradiction. Therefore, C is contained in the closed half-space determined by H that does not contain a, as claimed. $\square$

Lemma 3.13 *(Farkas Lemma, Version II) Given any $d \times n$ real matrix, A, and any vector, $z \in \mathbb{R}^d$, exactly one of the following alternatives occurs:*

(a) The linear system, $Ax = z$, has a solution, x, such that $x \geq 0$, or

(b) There is some $c \in \mathbb{R}^d$ such that $c^\top z < 0$ and $c^\top A \geq 0$.

Proof. The proof is analogous to the proof of Lemma 3.11 except that it uses Proposition 3.12 instead of Corollary 3.10 and either $z \in \mathrm{cone}(A_1, \ldots, A_n)$ or $z \notin \mathrm{cone}(A_1, \ldots, A_n)$. $\square$

One can show that Farkas II implies Farkas I. Here is another version of Farkas Lemma having to do with a system of inequalities, $Ax \leq z$. Although, this version may seem weaker that Farkas II, it is actually equivalent to it!

Lemma 3.14 *(Farkas Lemma, Version III) Given any $d \times n$ real matrix, A, and any vector, $z \in \mathbb{R}^d$, exactly one of the following alternatives occurs:*

(a) The system of inequalities, $Ax \leq z$, has a solution, x, or

(b) There is some $c \in \mathbb{R}^d$ such that $c \geq 0$, $c^\top z < 0$ and $c^\top A = 0$.

Proof. We use two tricks from linear programming:

1. We convert the system of inequalities, $Ax \leq z$, into a system of equations by introducing a vector of "slack variables", $\gamma = (\gamma_1, \ldots, \gamma_d)$, where the system of equations is
$$(A, I)\begin{pmatrix} x \\ \gamma \end{pmatrix} = z,$$
with $\gamma \geq 0$.

2. We replace each "unconstrained variable", x_i, by $x_i = X_i - Y_i$, with $X_i, Y_i \geq 0$.

Then, the original system $Ax \leq z$ has a solution, x (unconstrained), iff the system of equations
$$(A, -A, I)\begin{pmatrix} X \\ Y \\ \gamma \end{pmatrix} = z$$
has a solution with $X, Y, \gamma \geq 0$. By Farkas II, this system has no solution iff there exists some $c \in \mathbb{R}^d$ with $c^\top z < 0$ and
$$c^\top (A, -A, I) \geq 0,$$
that is, $c^\top A \geq 0$, $-c^\top A \geq 0$, and $c \geq 0$. However, these four conditions reduce to $c^\top z < 0$, $c^\top A = 0$ and $c \geq 0$. $\square$

These versions of Farkas lemma are statements of the form $(P \vee Q) \wedge \neg(P \wedge Q)$, which is easily seen to be equivalent to $\neg P \equiv Q$, namely, the logical equivalence of $\neg P$ and Q. Therefore, Farkas-type lemmas can be interpreted as criteria for the unsolvablity of various kinds of systems of linear equations or systems of linear inequalities, in the form of a separation property.

For example, Farkas II (Lemma 3.13) says that a system of linear equations, $Ax = z$, does not have any solution, $x \geq 0$, iff there is some $c \in \mathbb{R}^d$ such that $c^\top z < 0$ and $c^\top A \geq 0$. This means that there is a hyperplane, H, of equation $c^\top y = 0$, such that the columns vectors, A_j, forming the matrix A all lie in the positive closed half space, H_+, but z lies in the interior of the other half space, H_-, determined by H. Therefore, z can't be in the cone spanned by the A_j's.

Farkas III says that a system of linear inequalities, $Ax \leq z$, does not have any solution (at all) iff there is some $c \in \mathbb{R}^d$ such that $c \geq 0$, $c^\top z < 0$ and $c^\top A = 0$. This time, there is also a hyperplane of equation $c^\top y = 0$, with $c \geq 0$, such that the columns vectors, A_j, forming the matrix A all lie in H but z lies in the interior of the half space, H_-, determined by H. In the "easy" direction, if there is such a vector c and some x satisfying $Ax \leq b$, since $c \geq 0$, we get $c^\top Ax \leq x^\top z$, but $c^\top Ax = 0$ and $x^\top z < 0$, a a contradiction.

What is the crirerion for the insolvability of a system of inequalities $Ax \leq z$ with $x \geq 0$? This problem is equivalent to the insolvability of the set of inequalities
$$\begin{pmatrix} A \\ -I \end{pmatrix} x \leq \begin{pmatrix} z \\ 0 \end{pmatrix}$$

and by Farkas III, this system has no solution iff there is some vector, (c_1, c_2), with $(c_1, c_2) \geq 0$,

$$(c_1^\top, c_2^\top)\begin{pmatrix} A \\ -I \end{pmatrix} = 0 \quad \text{and} \quad (c_1^\top, c_2^\top)\begin{pmatrix} z \\ 0 \end{pmatrix} < 0.$$

The above conditions are equivalent to $c_1 \geq 0$, $c_2 \geq 0$, $c_1^\top A - c_2^\top = 0$ and $c_1^\top z < 0$, which reduce to $c_1 \geq 0$, $c_1^\top A \geq 0$ and $c_1^\top z < 0$.

We can put all these versions together to prove the following version of Farkas lemma:

Lemma 3.15 *(Farkas Lemma, Version IIIb) For any $d \times n$ real matrix, A, and any vector, $z \in \mathbb{R}^d$, the following statements are equivalent:*

(1) The system, $Ax = z$, has no solution $x \geq 0$ iff there is some $c \in \mathbb{R}^d$ such that $c^\top A \geq 0$ and $c^\top z < 0$.

(2) The system, $Ax \leq z$, has no solution iff there is some $c \in \mathbb{R}^d$ such that $c \geq 0$, $c^\top A = 0$ and $c^\top z < 0$.

(3) The system, $Ax \leq z$, has no solution $x \geq 0$ iff there is some $c \in \mathbb{R}^d$ such that $c \geq 0$, $c^\top A \geq 0$ and $c^\top z < 0$.

Proof. We already proved that (1) implies (2) and that (2) implies (3). The proof that (3) implies (1) is left as an easy exercise. □

The reader might wonder what is the criterion for the unsolvability of a system $Ax = z$, without any condition on x. However, since the unsolvability of the system $Ax = b$ is equivalent to the unsolvability of the system

$$\begin{pmatrix} A \\ -A \end{pmatrix} x \leq \begin{pmatrix} z \\ -z \end{pmatrix},$$

using (2), the above system is unsolvable iff there is some $(c_1, c_2) \geq (0, 0)$ such that

$$(c_1^\top, c_2^\top)\begin{pmatrix} A \\ -A \end{pmatrix} = 0 \quad \text{and} \quad (c_1^\top, c_2^\top)\begin{pmatrix} z \\ -z \end{pmatrix} < 0,$$

and these are equivalent to $c_1^\top A - c_2^\top A = 0$ and $c_1^\top z - c_2^\top z < 0$, namely, $c^\top A = 0$ and $c^\top z < 0$ where $c = c_1 - c_2 \in \mathbb{R}^d$. However, this simply says that the columns, $A_1, \ldots, A_n$, of A are linearly dependent and that z does not belong to the subspace spanned by $A_1, \ldots, A_n$, a criterion which we already knew from linear algebra.

As in Matousek and Gartner [28], we can summarize these various criteria in the following table:

	The system $Ax \leq z$	The system $Ax = z$
has no solution $x \geq 0$ iff	$\exists c \in \mathbb{R}^d$, such that $c \geq 0$, $c^\top A \geq 0$ and $c^\top z < 0$	$\exists c \in \mathbb{R}^d$, such that $c^\top A \geq 0$ and $c^\top z < 0$
has no solution $x \in \mathbb{R}^n$ iff	$\exists c \in \mathbb{R}^d$, such that, $c \geq 0$, $c^\top A = 0$ and $c^\top z < 0$	$\exists c \in \mathbb{R}^d$, such that $c^\top A = 0$ and $c^\top z < 0$

Remark: The strong duality theorem in linear programming can be proved using Lemma 3.15(c).

Finally, we have the separation theorem announced earlier for arbitrary nonempty convex subsets.

Theorem 3.16 *(Separation of disjoint convex sets) Given an affine space, X, let A and B be two nonempty disjoint convex subsets. Then, there is a hyperplane separating A and B.*

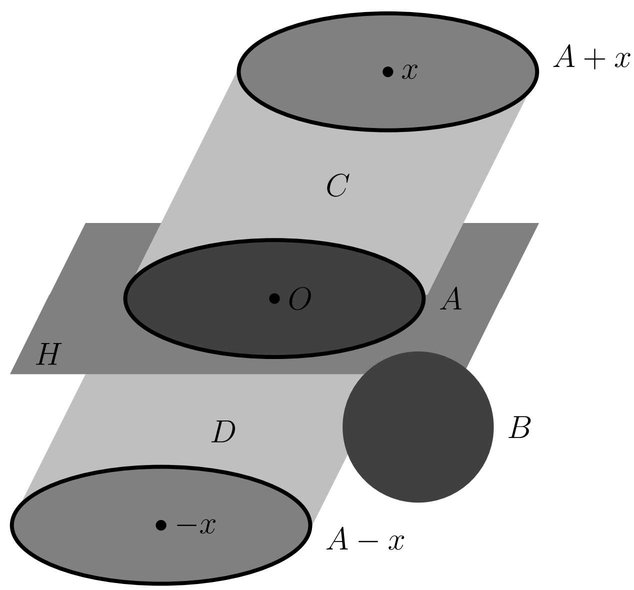


Figure 3.7: Separation Theorem, final version (Theorem 3.16)

Proof. The proof is by descending induction on $n = \dim A$. If $\dim A = \dim X$, we know from Proposition 3.3 that A has nonempty interior and we conclude using Corollary 3.8. Next, asssume that the induction hypothesis holds if $\dim A \geq n$ and assume $\dim A = n-1$. Pick an origin $O \in A$ and let H be a hyperplane containing A. Pick $x \in X$ outside H and define $C = \mathrm{conv}(A \cup \{A+x\})$ where $A+x = \{a+x \mid a \in A\}$ and $D = \mathrm{conv}(A \cup \{A-x\})$

where $A - x = \{a - x \mid a \in A\}$. Note that $C \cup D$ is convex. If $B \cap C \neq \emptyset$ and $B \cap D \neq \emptyset$, then the convexity of B and $C \cup D$ implies that $A \cap B \neq \emptyset$, a contradiction. Without loss of generality, assume that $B \cap C = \emptyset$. Since x is outside H, we have $\dim C = n$ and by the induction hypothesis, there is a hyperplane, H_1 separating C and B. As $A \subseteq C$, we see that H_1 also separates A and B. $\square$

Remarks:

(1) The reader should compare this proof (from Valentine [43], Chapter II) with Berger's proof using compactness of the projective space $\mathbb{P}^d$ [6] (Corollary 11.4.7).

(2) Rather than using the Hahn-Banach theorem to deduce separation results, one may proceed differently and use the following intuitively obvious lemma, as in Valentine [43] (Theorem 2.4):

Lemma 3.17 *If A and B are two nonempty convex sets such that $A \cup B = X$ and $A \cap B = \emptyset$, then $V = \overline{A} \cap \overline{B}$ is a hyperplane.*

One can then deduce Corollaries 3.8 and Theorem 3.16. Yet another approach is followed in Barvinok [3].

(3) How can some of the above results be generalized to infinite dimensional affine spaces, especially Theorem 3.6 and Corollary 3.8? One approach is to simultaneously relax the notion of interior and tighten a little the notion of closure, in a more "linear and less topological" fashion, as in Valentine [43].

Given any subset $A \subseteq X$ (where X may be infinite dimensional, but is a Hausdorff topological vector space), say that a point $x \in X$ is *linearly accessible from A* iff there is some $a \in A$ with $a \neq x$ and $]a, x[\subseteq A$. We let $\mathrm{lina}\, A$ be the set of all points linearly accessible from A and $\mathrm{lin}\, A = A \cup \mathrm{lina}\, A$.

A point $a \in A$ is a *core point of A* iff for every $y \in X$, with $y \neq a$, there is some $z \in]a, y[$, such that $[a, z] \subseteq A$. The set of all core points is denoted $\mathrm{core}\, A$.

It is not difficult to prove that $\mathrm{lin}\, A \subseteq \overline{A}$ and $\overset{\circ}{A} \subseteq \mathrm{core}\, A$. If A has nonempty interior, then $\mathrm{lin}\, A = \overline{A}$ and $\overset{\circ}{A} = \mathrm{core}\, A$. Also, if A is convex, then $\mathrm{core}\, A$ and $\mathrm{lin}\, A$ are convex. Then, Lemma 3.17 still holds (where X is not necessarily finite dimensional) if we redefine V as $V = \mathrm{lin}\, A \cap \mathrm{lin}\, B$ and allow the possibility that V could be X itself. Corollary 3.8 also holds in the general case if we assume that $\mathrm{core}\, A$ is nonempty. For details, see Valentine [43], Chapter I and II.

(4) Yet another approach is to define the notion of an algebraically open convex set, as in Barvinok [3]. A convex set, A, is *algebraically open* iff the intersection of A with every line, L, is an open interval, possibly empty or infinite at either end (or all of

L). An open convex set is algebraically open. Then, the Hahn-Banach theorem holds provided that A is an algebraically open convex set and similarly, Corollary 3.8 also holds provided A is algebraically open. For details, see Barvinok [3], Chapter 2 and 3. We do not know how the notion "algebraically open" relates to the concept of core.

(5) Theorems 3.6, 3.7 and Corollary 3.8 are proved in Lax [26] using the notion of *gauge function* in the more general case where A has some core point (but beware that Lax uses the terminology *interior point* instead of core point!).

An important special case of separation is the case where A is convex and $B = \{a\}$, for some point, a, in A.

3.2 Supporting Hyperplanes and Minkowski's Proposition

Recall the definition of a supporting hyperplane given in Definition 2.4. We have the following important proposition first proved by Minkowski (1896):

Proposition 3.18 *(Minkowski) Let A be a nonempty, closed, and convex subset. Then, for every point $a \in \partial A$, there is a supporting hyperplane to A through a.*

Proof. Let $d = \dim A$. If $d < \dim X$ (i.e., A has empty interior), then A is contained in some affine subspace V of dimension $d < \dim X$, and any hyperplane containing V is a supporting hyperplane for every $a \in A$. Now, assume $d = \dim X$, so that $\mathring{A} \neq \emptyset$. If $a \in \partial A$, then $\{a\} \cap \mathring{A} = \emptyset$. By Theorem 3.6, there is a hyperplane H separating $\mathring{A}$ and $L = \{a\}$. However, by Corollary 3.2, since $\mathring{A} \neq \emptyset$ and A is closed, we have

$$A = \overline{A} = \overline{\mathring{A}}.$$

Now, the half–space containing $\mathring{A}$ is closed, and thus, it contains $\overline{\mathring{A}} = A$. Therefore, H separates A and $\{a\}$. $\square$

Remark: The assumption that A is closed is convenient but unnecessary. Indeed, the proof of Proposition 3.18 shows that the proposition holds for every boundary point, $a \in \partial A$ (assuming $\partial A \neq \emptyset$).

Beware that Proposition 3.18 is false when the dimension of X is infinite and when $\mathring{A} = \emptyset$.

The proposition below gives a sufficient condition for a closed subset to be convex.

Proposition 3.19 *Let A be a closed subset with nonempty interior. If there is a supporting hyperplane for every point $a \in \partial A$, then A is convex.*

Proof. We leave it as an exercise (see Berger [6], Proposition 11.5.4). $\square$

The condition that A has nonempty interior is crucial!

The proposition below characterizes closed convex sets in terms of (closed) half–spaces. It is another intuitive fact whose rigorous proof is nontrivial.

Proposition 3.20 *Let A be a nonempty closed and convex subset. Then, A is the intersection of all the closed half–spaces containing it.*

Proof. Let A' be the intersection of all the closed half–spaces containing A. It is immediately checked that A' is closed and convex and that $A \subseteq A'$. Assume that $A' \neq A$, and pick $a \in A' - A$. Then, we can apply Corollary 3.10 to $\{a\}$ and A and we find a hyperplane, H, strictly separating A and $\{a\}$; this shows that A belongs to one of the two half-spaces determined by H, yet a does not belong to the same half-space, contradicting the definition of A'. $\square$

3.3 Polarity and Duality

Let $E = \mathbb{E}^n$ be a Euclidean space of dimension n. Pick any origin, O, in $\mathbb{E}^n$ (we may assume $O = (0, \ldots, 0)$). We know that the inner product on $E = \mathbb{E}^n$ induces a duality between E and its dual E^* (for example, see Chapter 6, Section 2 of Gallier [20]), namely, $u \mapsto \varphi_u$, where φ_u is the linear form defined by $\varphi_u(v) = u \cdot v$, for all $v \in E$. For geometric purposes, it is more convenient to recast this duality as a correspondence between points and hyperplanes, using the notion of polarity with respect to the unit sphere, $S^{n-1} = \{a \in \mathbb{E}^n \mid \|\mathbf{Oa}\| = 1\}$.

First, we need the following simple fact: For every hyperplane, H, not passing through O, there is a *unique* point, h, so that

$$H = \{a \in \mathbb{E}^n \mid \mathbf{Oh} \cdot \mathbf{Oa} = 1\}.$$

Indeed, any hyperplane, H, in $\mathbb{E}^n$ is the null set of some equation of the form

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta,$$

and if $O \notin H$, then $\beta \neq 0$. Thus, any hyperplane, H, not passing through O is defined by an equation of the form

$$h_1 x_1 + \cdots + h_n x_n = 1,$$

if we set $h_i = \alpha_i/\beta$. So, if we let $h = (h_1, \ldots, h_n)$, we see that

$$H = \{a \in \mathbb{E}^n \mid \mathbf{Oh} \cdot \mathbf{Oa} = 1\},$$

as claimed. Now, assume that

$$H = \{a \in \mathbb{E}^n \mid \mathbf{Oh_1} \cdot \mathbf{Oa} = 1\} = \{a \in \mathbb{E}^n \mid \mathbf{Oh_2} \cdot \mathbf{Oa} = 1\}.$$

The functions $a \mapsto \mathbf{Oh_1} \cdot \mathbf{Oa} - 1$ and $a \mapsto \mathbf{Oh_2} \cdot \mathbf{Oa} - 1$ are two affine forms defining the same hyperplane, so there is a nonzero scalar, λ, so that

$$\mathbf{Oh_1} \cdot \mathbf{Oa} - 1 = \lambda(\mathbf{Oh_2} \cdot \mathbf{Oa} - 1) \quad \text{for all } a \in \mathbb{E}^n$$

(see Gallier [20], Chapter 2, Section 2.10). In particular, for $a = O$, we find that $\lambda = 1$, and so,

$$\mathbf{Oh_1} \cdot \mathbf{Oa} = \mathbf{Oh_2} \cdot \mathbf{Oa} \quad \text{for all } a,$$

which implies $h_1 = h_2$. This proves the uniqueness of h.

Using the above, we make the following definition:

Definition 3.2 Given any point, $a \neq O$, the *polar hyperplane of a (w.r.t. S^{n-1})* or *dual of a* is the hyperplane, $a^\dagger$, given by

$$a^\dagger = \{b \in \mathbb{E}^n \mid \mathbf{Oa} \cdot \mathbf{Ob} = 1\}.$$

Given a hyperplane, H, not containing O, the *pole of H (w.r.t S^{n-1})* or *dual of H* is the (unique) point, $H^\dagger$, so that

$$H = \{a \in \mathbb{E}^n \mid \mathbf{OH}^\dagger \cdot \mathbf{Oa} = 1\}.$$

We often abbreviate polar hyperplane to polar. We immediately check that $a^{\dagger\dagger} = a$ and $H^{\dagger\dagger} = H$, so, we obtain a bijective correspondence between $\mathbb{E}^n - \{O\}$ and the set of hyperplanes not passing through O.

When a is outside the sphere S^{n-1}, there is a nice geometric interpetation for the polar hyperplane, $H = a^\dagger$. Indeed, in this case, since

$$H = a^\dagger = \{b \in \mathbb{E}^n \mid \mathbf{Oa} \cdot \mathbf{Ob} = 1\}$$

and $\|\mathbf{Oa}\| > 1$, the hyperplane H intersects S^{n-1} (along an $(n-2)$-dimensional sphere) and if b is any point on $H \cap S^{n-1}$, we claim that $\mathbf{Ob}$ and $\mathbf{ba}$ are orthogonal. This means that $H \cap S^{n-1}$ is the set of points on S^{n-1} where the lines through a and tangent to S^{n-1} touch S^{n-1} (they form a cone tangent to S^{n-1} with apex a). Indeed, as $\mathbf{Oa} = \mathbf{Ob} + \mathbf{ba}$ and $b \in H \cap S^{n-1}$ i.e., $\mathbf{Oa} \cdot \mathbf{Ob} = 1$ and $\|\mathbf{Ob}\|^2 = 1$, we get

$$1 = \mathbf{Oa} \cdot \mathbf{Ob} = (\mathbf{Ob} + \mathbf{ba}) \cdot \mathbf{Ob} = \|\mathbf{Ob}\|^2 + \mathbf{ba} \cdot \mathbf{Ob} = 1 + \mathbf{ba} \cdot \mathbf{Ob},$$

which implies $\mathbf{ba} \cdot \mathbf{Ob} = 0$. When $a \in S^{n-1}$, the hyperplane $a^\dagger$ is tangent to S^{n-1} at a.

Also, observe that for any point, $a \neq O$, and any hyperplane, H, not passing through O, if $a \in H$, then, $H^\dagger \in a^\dagger$, i.e, the pole, $H^\dagger$, of H belongs to the polar, $a^\dagger$, of a. Indeed, $H^\dagger$ is the unique point so that

$$H = \{b \in \mathbb{E}^n \mid \mathbf{OH}^\dagger \cdot \mathbf{Ob} = 1\}$$

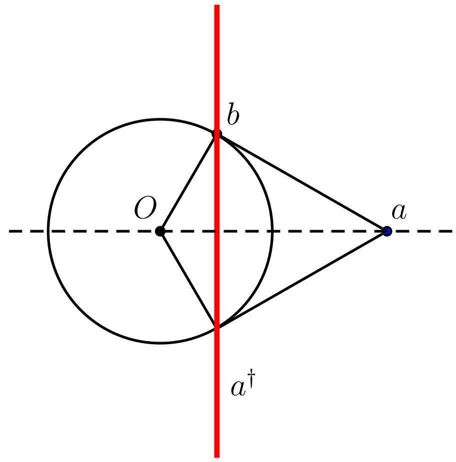


Figure 3.8: The polar, $a^\dagger$, of a point, a, outside the sphere S^{n-1}

and

$$a^\dagger = \{b \in \mathbb{E}^n \mid \mathbf{Oa} \cdot \mathbf{Ob} = 1\};$$

since $a \in H$, we have $\mathbf{OH}^\dagger \cdot \mathbf{Oa} = 1$, which shows that $H^\dagger \in a^\dagger$.

If $a = (a_1, \ldots, a_n)$, the equation of the polar hyperplane, $a^\dagger$, is

$$a_1 X_1 + \cdots + a_n X_n = 1.$$

Remark: As we noted, polarity in a Euclidean space suffers from the minor defect that the polar of the origin is undefined and, similarly, the pole of a hyperplane through the origin does not make sense. If we embed $\mathbb{E}^n$ into the projective space, $\mathbb{P}^n$, by adding a "hyperplane at infinity" (a copy of $\mathbb{P}^{n-1}$), thereby viewing $\mathbb{P}^n$ as the disjoint union $\mathbb{P}^n = \mathbb{E}^n \cup \mathbb{P}^{n-1}$, then the polarity correspondence can be defined everywhere. Indeed, the polar of the origin is the hyperplane at infinity ($\mathbb{P}^{n-1}$) and since $\mathbb{P}^{n-1}$ can be viewed as the set of hyperplanes through the origin in $\mathbb{E}^n$, the pole of a hyperplane through the origin is the corresponding "point at infinity" in $\mathbb{P}^{n-1}$.

Now, we would like to extend this correspondence to subsets of $\mathbb{E}^n$, in particular, to convex sets. Given a hyperplane, H, not containing O, we denote by H_- the closed half-space containing O.

Definition 3.3 Given any subset, A, of $\mathbb{E}^n$, the set

$$A^* = \{b \in \mathbb{E}^n \mid \mathbf{Oa} \cdot \mathbf{Ob} \leq 1, \quad \text{for all } a \in A\} = \bigcap_{\substack{a \in A \\ a \neq O}} (a^\dagger)_-,$$

is called the *polar dual* or *reciprocal* of A.

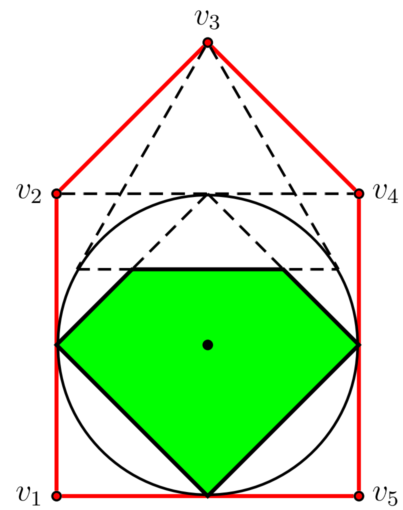


Figure 3.9: The polar dual of a polygon

For simplicity of notation, we write $a_-^\dagger$ for $(a^\dagger)_-$. Observe that $\{O\}^* = \mathbb{E}^n$, so it is convenient to set $O_-^\dagger = \mathbb{E}^n$, even though $O^\dagger$ is undefined. By definition, A^* is convex even if A is not. Furthermore, note that

(1) $A \subseteq A^{**}$.

(2) If $A \subseteq B$, then $B^* \subseteq A^*$.

(3) If A is convex and closed, then $A^* = (\partial A)^*$.

It follows immediately from (1) and (2) that $A^{***} = A^*$. Also, if $B^n(r)$ is the (closed) ball of radius $r > 0$ and center O, it is obvious by definition that $B^n(r)^* = B^n(1/r)$.

In Figure 3.9, the polar dual of the polygon $(v_1, v_2, v_3, v_4, v_5)$ is the polygon shown in green. This polygon is cut out by the half-planes determined by the polars of the vertices $(v_1, v_2, v_3, v_4, v_5)$ and containing the center of the circle. These polar lines are all easy to determine by drawing for each vertex, v_i, the tangent lines to the circle and joining the contact points. The construction of the polar of v_3 is shown in detail.

Remark: We chose a different notation for polar hyperplanes and polars ($a^\dagger$ and $H^\dagger$) and polar duals (A^*), to avoid the potential confusion between $H^\dagger$ and H^*, where H is a hyperplane (or $a^\dagger$ and $\{a\}^*$, where a is a point). Indeed, they are completely different! For example, the polar dual of a hyperplane is either a line orthogonal to H through O, if $O \in H$, or a semi-infinite line through O and orthogonal to H whose endpoint is the pole, $H^\dagger$, of H, whereas, $H^\dagger$ is a single point! Ziegler ([45], Chapter 2) use the notation $A^\triangle$ instead of A^* for the polar dual of A.

We would like to investigate the duality induced by the operation $A \mapsto A^*$. Unfortunately, it is not always the case that $A^{**} = A$, but this is true when A is closed and convex, as shown in the following proposition:

Proposition 3.21 *Let A be any subset of $\mathbb{E}^n$ (with origin O).*

(i) If A is bounded, then $O \in \overset{\circ}{A^*}$; if $O \in \overset{\circ}{A}$, then A^* is bounded.

(ii) If A is a closed and convex subset containing O, then $A^{**} = A$.

Proof. (i) If A is bounded, then $A \subseteq B^n(r)$ for some $r > 0$ large enough. Then, $B^n(r)^* = B^n(1/r) \subseteq A^*$, so that $O \in \overset{\circ}{A^*}$. If $O \in \overset{\circ}{A}$, then $B^n(r) \subseteq A$ for some r small enough, so $A^* \subseteq B^n(r)^* = B^r(1/r)$ and A^* is bounded.

(ii) We always have $A \subseteq A^{**}$. We prove that if $b \notin A$, then $b \notin A^{**}$; this shows that $A^{**} \subseteq A$ and thus, $A = A^{**}$. Since A is closed and convex and $\{b\}$ is compact (and convex!), by Corollary 3.10, there is a hyperplane, H, strictly separating A and b and, in particular, $O \notin H$, as $O \in A$. If $h = H^\dagger$ is the pole of H, we have

$$\mathbf{Oh} \cdot \mathbf{Ob} > 1 \quad \text{and} \quad \mathbf{Oh} \cdot \mathbf{Oa} < 1, \quad \text{for all } a \in A$$

since $H_- = \{a \in \mathbb{E}^n \mid \mathbf{Oh} \cdot \mathbf{Oa} \leq 1\}$. This shows that $b \notin A^{**}$, since

$$\begin{aligned} A^{**} &= \{c \in \mathbb{E}^n \mid \mathbf{Od} \cdot \mathbf{Oc} \leq 1 \quad \text{for all } d \in A^*\} \\ &= \{c \in \mathbb{E}^n \mid (\forall d \in \mathbb{E}^n)(\text{if} \quad \mathbf{Od} \cdot \mathbf{Oa} \leq 1 \quad \text{for all } a \in A, \quad \text{then} \quad \mathbf{Od} \cdot \mathbf{Oc} \leq 1)\}, \end{aligned}$$

just let $c = b$ and $d = h$. $\square$

Remark: For an arbitrary subset, $A \subseteq \mathbb{E}^n$, it can be shown that $A^{**} = \overline{\mathrm{conv}(A \cup \{O\})}$, the topological closure of the convex hull of $A \cup \{O\}$.

Proposition 3.21 will play a key role in studying polytopes, but before doing this, we need one more proposition.

Proposition 3.22 *Let A be any closed convex subset of $\mathbb{E}^n$ such that $O \in \overset{\circ}{A}$. The polar hyperplanes of the points of the boundary of A constitute the set of supporting hyperplanes of A^*. Furthermore, for any $a \in \partial A$, the points of A^* where $H = a^\dagger$ is a supporting hyperplane of A^* are the poles of supporting hyperplanes of A at a.*

Proof. Since $O \in \overset{\circ}{A}$, we have $O \notin \partial A$, and so, for every $a \in \partial A$, the polar hyperplane $a^\dagger$ is well-defined. Pick any $a \in \partial A$ and let $H = a^\dagger$ be its polar hyperplane. By definition, $A^* \subseteq H_-$, the closed half-space determined by H and containing O. If T is any supporting hyperplane to A at a, as $a \in T$, we have $t = T^\dagger \in a^\dagger = H$. Furthermore, it is a simple exercise to prove that $t \in (T_-)^*$ (in fact, $(T_-)^*$ is the interval with endpoints O and t). Since $A \subseteq T_-$ (because T is a supporting hyperplane to A at a), we deduce that $t \in A^*$, and thus, H is a supporting hyperplane to A^* at t. By Proposition 3.21, as A is closed and convex, $A^{**} = A$; it follows that all supporting hyperplanes to A^* are indeed obtained this way. $\square$