Chapter 3

Basic Properties of Convex Sets

3.1 Convex Sets

Convex sets play a very important role in geometry. In this chapter, we state 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 $0 \leq \lambda \leq 1$ ($\lambda \in \mathbb{R}$).

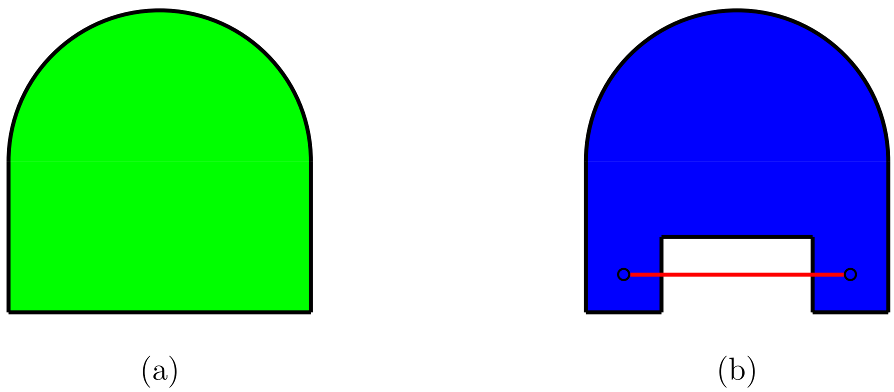

Figure 3.1: (a) A convex set; (b) A nonconvex set

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.

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 $\text{conv}(S)$) 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 is denoted by $\langle S \rangle$ or $\text{aff}(S)$.

Definition 3.1.1 The *dimension* of a nonempty convex subset, S, of X, denoted by $\dim S$, is the dimension of the smallest affine subset $\langle S \rangle$ containing S.

Lemma 3.1.2 *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 $(a_i)_{i \in I}$.*

In view of lemma 3.1.2, it is obvious that any affine subspace of E is convex.

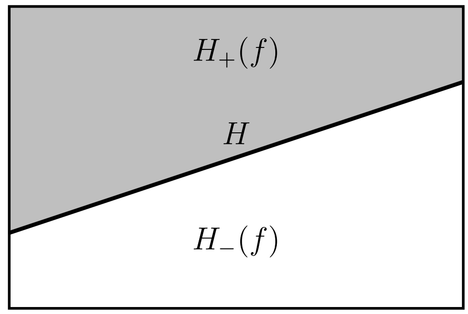

Figure 3.2: The two half-spaces determined by a hyperplane, H

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., $H = \operatorname{Ker} f$), we can define the two subsets

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

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

$$H_+(f) \cup H_-(f) = E \quad \text{and} \quad H_+(f) \cap H_-(f) = H.$$

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*).

It is natural to wonder whether lemma 3.1.2 can be sharpened in two directions:

(1) is it possible 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 only consider 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.

In case 2, the theorem of *Krein and Milman* asserts that a convex set which is also compact is the convex hull of its extremal points (see Berger [**?**] or Lang [**?**]).

First, we will prove *Carathéodory's Theorem*.

3.2 Carathéodory's Theorem

The following technical (and dull!) lemma plays a crucial role in the proof.

Lemma 3.2.1 *Given an affine space* $\langle E, \overrightarrow{E}, +\rangle$, *let* $(a_i)_{i\in I}$ *be a family of points in* E. *The family* $(a_i)_{i\in I}$ *is affinely dependent iff there is a family* $(\lambda_i)_{i\in I}$ *such that* $\lambda_j \neq 0$ *for some* $j \in I$, $\sum_{i\in I} \lambda_i = 0$, *and* $\sum_{i\in I} \lambda_i \mathbf{xa_i} = 0$ *for every* $x \in E$.

Theorem 3.2.2 *Given any affine space* E *of dimension* m, *for any (nonempty) 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 sketch. By lemma 3.1.2,

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

We would like to prove that

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

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 less than q points in S (with $|J| < q$).

We shall prove that b can be written as a convex combination of $q-1$ of the a_i. Since E has dimension m and $q \geq m+2$, the points $a_1, \ldots, a_q$ must be affinely dependent, and we use lemma 3.2.1. □

If S is a finite (of infinite) set of points in the affine plane $\mathbb{A}^2$, theorem 3.2.2 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 3.2.3 *If K is any compact subset of $\mathbb{A}^m$, then the convex hull,* $\mathrm{conv}(K)$, *of K is also compact.*

There is also a version of Theorem 3.2.2 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 3.2.4 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 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, cone(S), is called a *polyhedral cone*.

Theorem 3.2.5 *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.*

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 3.4).

Theorem 3.2.6 *(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 3.2.6 is not hard to prove, we will not prove it here. Instead, we refer the reader to Matousek [**?**], Chapter 8, Section 8.2.

There is also a stronger version of Theorem 3.2.6, in which it is enough to assume that $b \in \mathrm{conv}(S_i \cup S_j)$ for all i, j with $1 \leq i < j \leq m+1$.

3.3 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

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

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

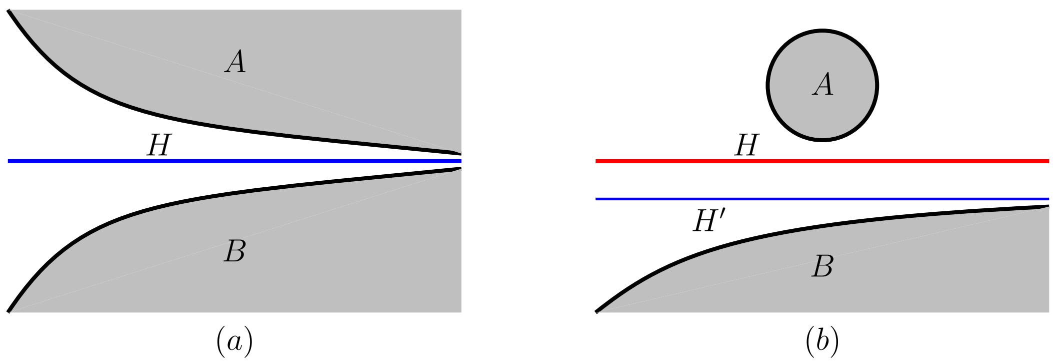

Figure 3.3: (a) A separating hyperplane, H. (b) A strictly separating hyperplane, H

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

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

The set $\{\overset{\circ}{H}_+(f), \overset{\circ}{H}_-(f)\}$ only depends on the hyperplane H.

Clearly, $\overset{\circ}{H}_+(f) = H_+(f) - H$ and $\overset{\circ}{H}_-(f) = H_-(f) - H$.

Definition 3.3.1 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.

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

Definition 3.3.2 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 $a \in \partial A$.

Also, if A is convex, then $H \cap \overset{\circ}{A} = \emptyset$.

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 3.4).

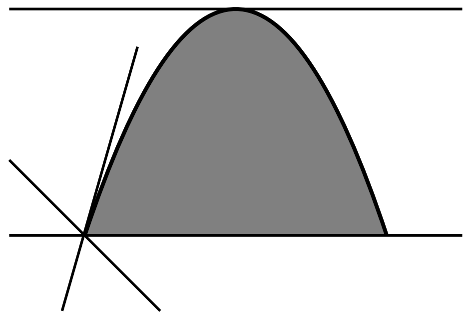

Figure 3.4: Examples of supporting hyperplanes

Next, we consider various types of boundary points of closed convex sets.

Definition 3.3.3 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.

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 [**?**] (Proposition 11.6.2):

Proposition 3.3.4 *The set of vertices of a closed and convex subset is countable.*

Another important concept is that of an extremal point.

Definition 3.3.5 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 $[x, y] \subseteq A$ ($x \neq y$, $a \neq x$, $a \neq y$).

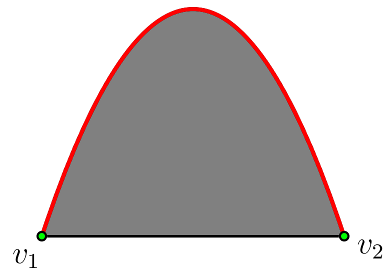

Figure 3.5: Examples of vertices and extreme points

Observe that a vertex is extremal, but the converse is false. For example, in Figure 3.5, 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 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.

Lemma 3.3.6 *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.*

The following important theorem shows that only extremal points matter as far as determining a compact and convex subset from its boundary.

The proof uses a proposition due to Minkowski (Proposition 4.2.1) which will be proved in the next chapter.

Theorem 3.3.7 *(Krein and Milman) 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.*

Observe that Krein and Milman's theorem implies that any nonemty 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 holds for infinite dimensional affine spaces, provided that they are locally convex.

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 3.3.8 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

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

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 3.3.9 *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.*

Proposition 3.3.9 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 3.3.9 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.

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

We begin with *Radon's theorem*.

Theorem 3.4.1 *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.*

A partition, (X_1, X_2), of X satisfying the conditions of Theorem 3.4.1 is sometimes called a *Radon partition* of X. A point in $\mathrm{conv}(X_1) \cap \mathrm{conv}(X_2)$ is called a *Radon point*. Figure 3.6 shows two Radon partitions of five points in the plane.

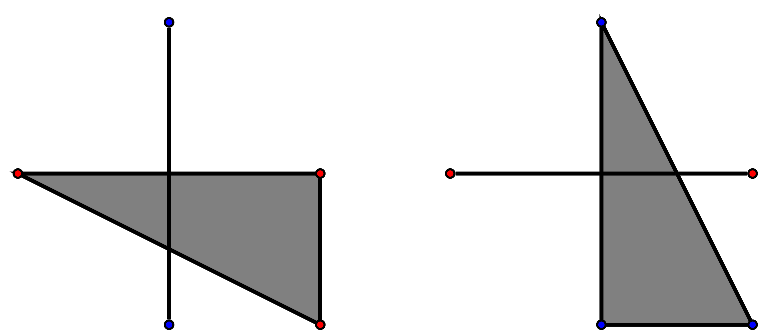

Figure 3.6: Examples of Radon Partitions

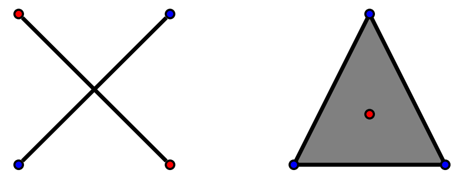

Figure 3.7: The Radon Partitions of four points (in $\mathbb{A}^2$)

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 3.7.

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.

Theorem 3.4.2 *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* $\{0\}$.

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

Theorem 3.4.3 *(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 3.4.3 is called a *Tverberg partition* and a point in $\bigcap_{i=1}^{r} \mathrm{conv}(X_i)$ is called a *Tverberg point*.

Theorem 3.4.3 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 [**?**], Chapter 8, Section 8.3. A *colored Tverberg theorem* and more can also be found in Matousek [**?**] (Section 8.3).

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

Theorem 3.4.4 *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.*

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.

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 $n/2$).

Definition 3.4.5 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 $\frac{n}{d+1}$).

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.*

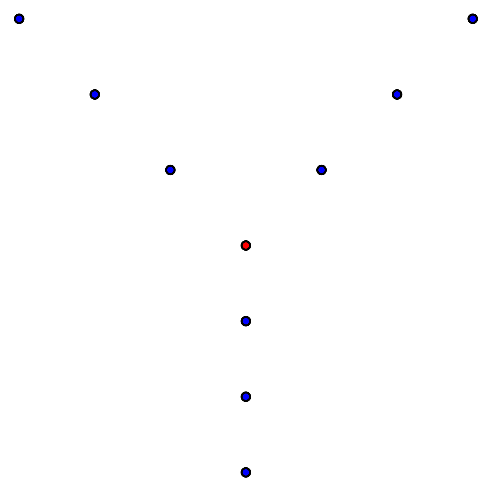

Figure 3.8: Example of a centerpoint

Example 3.8 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, $\mathring{H}_+$ (resp. $\mathring{H}_-$) containing at least $\frac{dn}{d+1} + 1$ points from S (again, we mean $\lceil \frac{dn}{d+1} \rceil + 1$).

We are now ready to prove the existence of centerpoints.

Theorem 3.4.6 *Every finite set, $S = \{a_1, \ldots, a_n\}$, of n points in $\mathbb{A}^d$ has some centerpoint.*

The proof is by induction and its uses the second characterization of centerpoints involving open half-spaces containing at least $\frac{dn}{d+1} + 1$ points.

The proof actually shows that the set of centerpoints of S is a convex set.

It should also be noted that Theorem 3.4.6 can be proved easily using Tverberg's theorem (Theorem 3.4.3). Indeed, for a judicious choice of r, any Tverberg point is a centerpoint!

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*.

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).

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 4

Separation and Supporting Hyperplanes

4.1 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 $\mathring{A}$, and is called the *interior of A*.

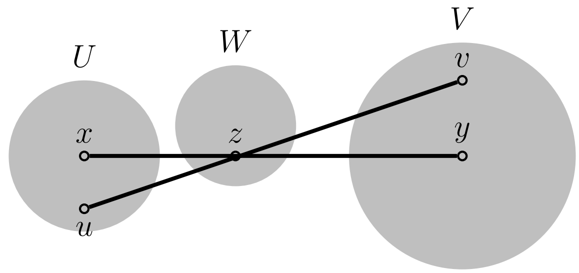

Figure 4.1: Illustration for the proof of Lemma 4.1.1

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

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

Lemma 4.1.1 *Let S be a nonempty convex set, and let $x \in \overset{\circ}{S}$ and $y \in \overline{S}$. Then, we have $]x, y[\subseteq S$.*

Corollary 4.1.2 *If S is convex, then $\overset{\circ}{S}$ is also convex and we have $\overset{\circ}{S} = \overset{\circ}{\overline{S}}$. Further, if $\overset{\circ}{S} \neq \emptyset$, then $\overline{S} = \overline{\overset{\circ}{S}}$.*

Beware that if S is a closed set, then its convex hull, $\mathrm{conv}(S)$, is *not* necessarily closed! However, $\mathrm{conv}(S)$ is closed when S is compact (see Proposition 3.2.3).

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.

Proposition 4.1.3 *A nonempty convex set S has a nonempty interior iff $\dim S = \dim X$.*

Proposition 4.1.3 is false in infinite dimension.

Proposition 4.1.4 *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.

Definition 4.1.5 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 4.1.6 *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 $L \cap B = \emptyset$.*

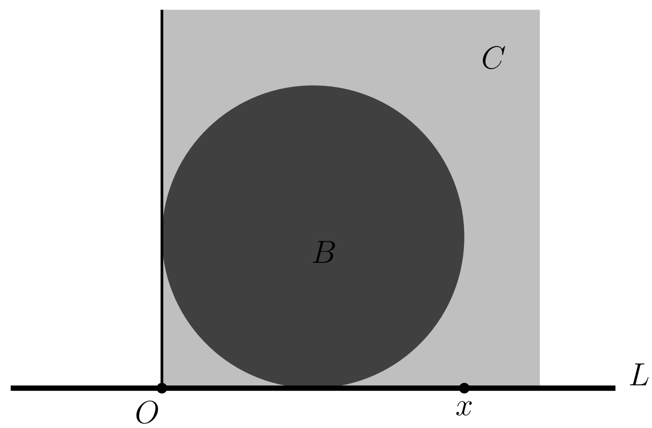

Figure 4.2: Hahn-Banach Theorem in the plane (Lemma 4.1.6)

Finally, we come to the Hahn-Banach theorem.

Theorem 4.1.7 *(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 $\dim X = 2$. $\square$

Remark: The geometric form of the Hahn-Banach theorem also holds when the dimension of X is infinite, but a more sophisticated proof is required (it uses Zorn's lemma).

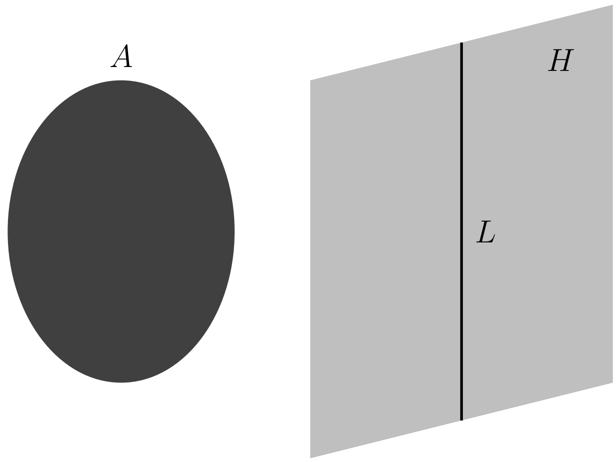

Figure 4.3: Hahn-Banach Theorem, geometric form (Theorem 4.1.7)

Theorem 4.1.7 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 close 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).

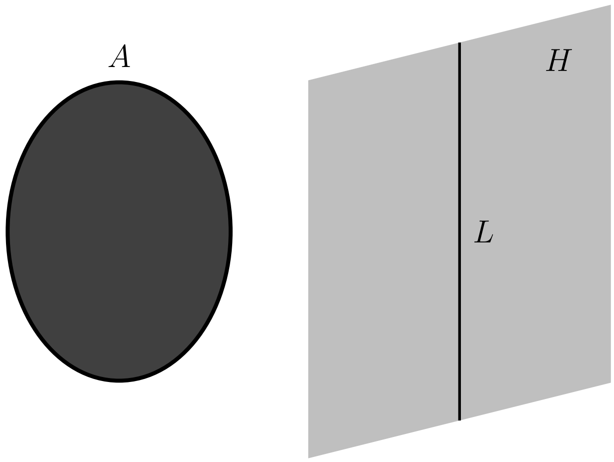

Figure 4.4: Hahn-Banach Theorem, second version (Theorem 4.1.8)

Theorem 4.1.7 has many important corollaries.

We begin with the following version of the Hahn-Banach theorem:

Theorem 4.1.8 *(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.

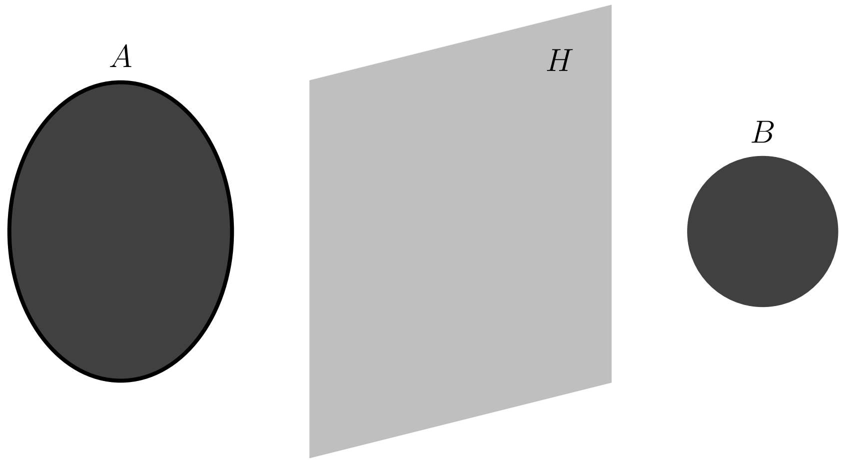

Figure 4.5: Separation Theorem, version 1 (Corollary 4.1.9)

Corollary 4.1.9 *Given an affine space, X, let A and B be two nonempty disjoint convex subsets and assume that A has nonempty interior ($\overset{\circ}{A} \neq \emptyset$). Then, there is a hyperplane separating A and B.*

Remark: Theorem 4.1.8 and Corollary 4.1.9 also hold in the infinite case.

Corollary 4.1.10 *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 4.1.10 *fails* for *closed* convex sets. However, Corollary 4.1.10 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

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

Corollary 4.1.11 *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.*

A "cute" application of Corollary 4.1.11 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 4.1.12 *(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 $c^\top A \geq \alpha$.

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,

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

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:

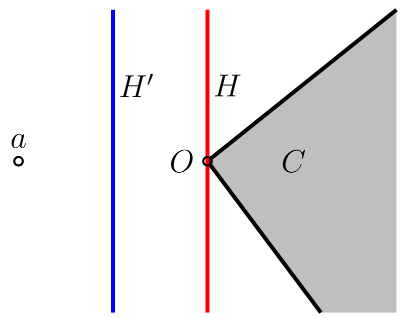

Figure 4.6: Illustration for the proof of Proposition 4.1.13

Proposition 4.1.13 *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 $a \notin H$.*

Lemma 4.1.14 *(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 $c^\top A \geq 0$.

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 4.1.15 *(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 $c^\top A = 0$.

The proof uses 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 $c \geq 0$. $\square$

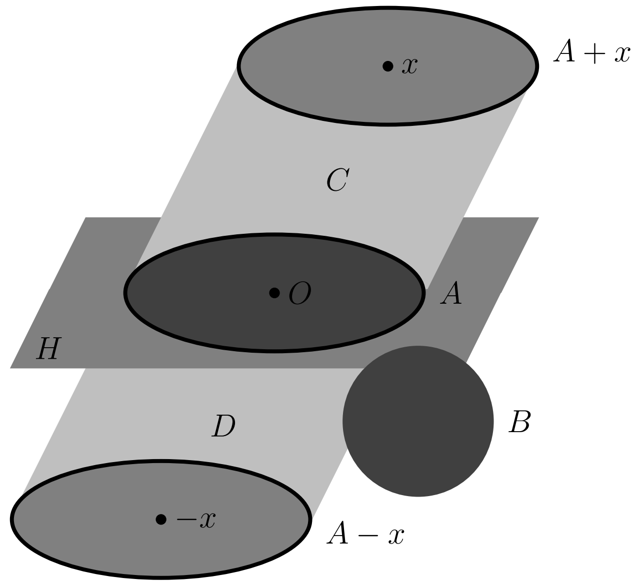

Figure 4.7: Separation Theorem, final version (Theorem 4.1.16)

Finally, we have the separation theorem announced earlier for arbitrary nonempty convex subsets. The proof is by descending induction on $\dim(A)$. (For a different proof, see Berger [**?**], Corollary 11.4.7.)

Corollary 4.1.16 *(Separation Theorem, final version)* *Given an affine space, X, let A and B be two nonempty disjoint convex subsets. Then, there is a hyperplane separating A and B.*

Remarks:

(1) The reader should compare the proof from Valentine [**?**], Chapter II with Berger's proof using compactness of the projective space $\mathbb{P}^d$ [**?**] (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 [**?**] (Theorem 2.4):

Lemma 4.1.17 *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 4.1.9 and 4.1.16. Yet another approach is followed in Barvinok [**?**].

(3) How can some of the above results be generalized to infinite dimensional affine spaces, especially Theorem 4.1.7 and Corollary 4.1.9?

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 [**?**].

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 core A and lin A are convex. Then, Lemma 4.1.17 still holds (where X is not necessarily finite dimensional) if we redefine V as $V = \text{lin}\, A \cap \text{lin}\, B$ and allow the possibility that V could be X itself.

Corollary 4.1.9 also holds in the general case if we assume that core A is nonempty. For details, see Valentine [**?**], Chapter I and II.

(4) Yet another approach is to define the notion of an algebraically open convex set, as in Barvinok [**?**].

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 4.1.9 also holds provided A is algebraically open.

For details, see Barvinok [**?**], Chapter 2 and 3. We do not know how the notion "algebraically open" relates to the concept of core.

(5) Theorems 4.1.7, 4.1.8 and Corollary 4.1.9 are proved in Lax 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.

4.2 Supporting Hyperplanes and Minkowski's Proposition

Recall the definition of a supporting hyperplane given in Definition 3.3.2. We have the following important proposition first proved by Minkowski (1896):

Proposition 4.2.1 *(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.*

Beware that Proposition 4.2.1 is false when the dimension X of A is infinite and when $\overset{\circ}{A} = \emptyset$.

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

Proposition 4.2.2 *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.*

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