Convex geometry

2023/24 spring

`BMETE94MM16` and `BMETE94AM22`

This document is a mildly edited version of the lecture notes by Zsolt Lángi (available at the course webpage for the 2022/23 spring term). Updated: April 14, 2024.

Contents

1 Affine subspace, affine combination

In this lecture we introduce the basic concepts used throughout the semester.

We deal with only finite dimensional Euclidean spaces. We regard an n-dimensional Euclidean spaces as an affine space whose vectors are the elements of the n-dimensional vector space $\mathbb{R}^n$ over the set of real numbers. Fixing an arbitrary point of an affine space, the elements of the corresponding vector space and the points of the space can be identified in a natural way, in which a point is associated to the vector that moves the fixed point to this one. In this case the fixed point is usually called *origin*. As it often appears in the literature, during the term we identify the Euclidean space with the vector space $\mathbb{R}^n$ (in high school language: we identify points and their position vectors). We will usually denote the points/vectors of the space $\mathbb{R}^n$ by small Latin letters, while its subsets by capital Latin letters.

We denote the usual inner (scalar) product of $\mathbb{R}^n$ by $\langle .,. \rangle$. The length $\|v\|$ of a vector $v \in \mathbb{R}^n$ is the quantity $\sqrt{\langle v, v \rangle}$. For the coordinates of the vector/point v in the standard orthonormal basis of $\mathbb{R}^n$ we use the notation $v = (v_1, v_2, \ldots, v_n)$. We denote the origin by o. The *distance* of the points $p = (x_1, x_2, \ldots, x_n)$ and $q = (x'_1, x'_2, \ldots, x'_n)$, denoted by $\operatorname{dist}(p, q)$, is the quantity $\sqrt{\sum_{i=1}^{n} (x'_i - x_i)^2}$, which coincides with the value of $\|q - p\|$. The *interior*, *boundary*, *closure* and *cardinality* of a set $X \subseteq \mathbb{R}^n$ will be denoted by $\operatorname{int}(X), \operatorname{bd}(X), \operatorname{cl}(X), |X|$, respectively.

Definition 1.1. Let V_1 and V_2 be two point sets, and $\lambda \in \mathbb{R}$. Then

$$V_1 + V_2 = \{v_1 + v_2 | v_1 \in V_1, v_2 \in V_2\}$$

is called the *Minkowski sum of the two sets*, and

$$\lambda V_1 = \{\lambda v_1 | v_1 \in V_1\}$$

the *multiple* of V_1 by λ.

Definition 1.2. Let $p \in \mathbb{R}^n$ be an arbitrary point, and L an arbitrary (linear) subspace in the vector space $\mathbb{R}^n$. Then the set $p + L \subseteq \mathbb{R}^n$ is called an *affine subspace* of the space $\mathbb{R}^n$.

The next remark is a straightforward consequence of the properties of linear subspaces.

Remark 1.3. *Let $p, q \in \mathbb{R}^n$ and let L, L' be linear subspaces in $\mathbb{R}^n$. Then $p + L = q + L'$ is satisfied if and only if $L = L'$ and $q \in p + L$.*

Proof. Assume that $p + L = q + L'$. Then $L = (q - p) + L'$ by the definition of Minkowski sum, which yields, in particular, that $q - p \in L$, from which we have $q \in p + L$. But as linear subspaces are closed with respect to addition, $q - p \in L$ implies $(q - p) + L = L$, from which $(q - p) + L = (q - p) + L'$, yielding $L = L'$. On the other hand, if $q \in p + L$, then $(q - p) \in L \Longrightarrow (q - p) + L = L \Longrightarrow q + L = p + L$. $\square$

Theorem 1.4. *A nonempty intersection of affine subspaces is an affine subspace.*

Proof. Consider the affine subspaces A_i $(i \in I)$, where I is an arbitrary index set. Let $A = \bigcap_{i \in I} A_i$. Consider a point $p \in A$. Then, due to the previous remark, for any $i \in I$ we have $A_i = p + L_i$ for some suitable linear subspace L_i of $\mathbb{R}^n$ (see Figure 1). The intersection of linear subspaces is a linear subspace, and thus, $L = \bigcap_{i \in I} L_i$ is a linear subspace. On the other hand, we clearly have $A = p + L$, from which the assertion follows. □

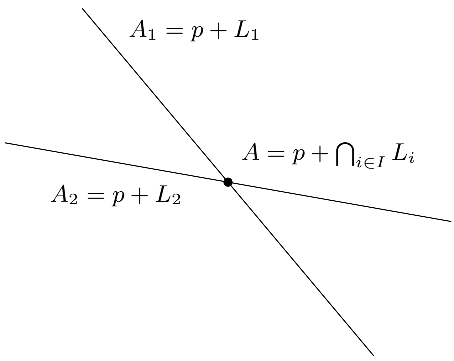


Figure 1: If p is a point in the intersection of affine subspaces A_i, then $A_i = p + L_i$ for suitable linear subspaces L_i and the intersection of the affine subspaces is the affine subspace $p + \bigcap_{i \in I} L_i$.

By the *dimension* of an affine subspace we mean the dimension of the corresponding linear subspace. We call the 0-, 1-, 2-, $(n-1)$-dimensional subspaces *points, lines, planes* and *hyperplanes*. A k-dimensional affine subspace may also be called a k-flat.

The next property readily follows from the definition of affine subspaces and the properties of the inner product.

Remark 1.5. *If $u \in \mathbb{R}^n$ and $t \in \mathbb{R}$ arbitrary, then the set $\{v \in \mathbb{R}^n | \langle v, u \rangle = t\}$ is a hyperplane (Figure 2). Furthermore, for any hyperplane H there is some vector $u \in \mathbb{R}^n$ and scalar $t \in \mathbb{R}$ for which $H = \{v \in \mathbb{R}^n | \langle v, u \rangle = t\}$.*

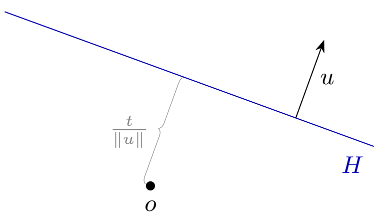


Figure 2: Hyperplanes are precisely sets of the form $H = \{v \in \mathbb{R}^n | \langle u, v \rangle = t\}$ with $u \neq 0$. The vector u is a normal vector of H and the distance of H and o is $|t| / \|u\|$. If u points away from the origin then $t > 0$, while in the opposite case $t < 0$. The hyperplane passes through the origin iff $t = 0$.

Since inner product is a continuous map from $\mathbb{R}^n$ to $\mathbb{R}$, the previous remark implies that for any hyperplane H decomposes the space into two connected, open components, which we call *open half spaces*. The unions of open half spaces with the bounding hyperplane we call *closed half spaces*.

Definition 1.6. Let $G_1 = p_1 + L_1$ and $G_2 = p_2 + L_2$ be affine subspaces. If for any vectors $v_1 \in L_1$, $v_2 \in L_2$ we have $\langle v_1, v_2 \rangle = 0$, then we say that G_1 and G_2 are *perpendicular* or *orthogonal*. Two affine subspaces are *parallel*, if they can be written in the form $p_1 + L$ and $p_2 + L$, where L is a linear subspace (Figure 3).

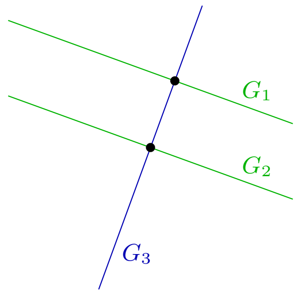


Figure 3: The affine subspaces G_1 and G_2 are parallel, and both are perpendicular to G_3.

Definition 1.7. Let $X \subset \mathbb{R}^n$ be a nonempty set. Then the *affine hull* of X, denoted by $\mathrm{aff}(X)$, is defined as the intersection of all affine subspaces containing X (Figure 4). The *linear hull* of X is defined as the affine hull $\mathrm{aff}(X \cup \{o\})$. We denote the linear hull of X by $\mathrm{lin}(X)$. The *relative interior* and *relative boundary* of X is defined as the interior and boundary of X, respectively, with respect to the induced topology in $\mathrm{aff}(X)$. We denote them by $\mathrm{relint}(X)$ and $\mathrm{relbd}(X)$, respectively (Figure 5).

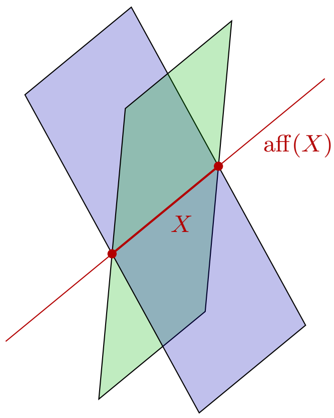


Figure 4: The affine hull of X is the intersection of all affine subspaces containing X.

We remark that by Theorem 1.4, the affine hull of a set is an affine subspace.

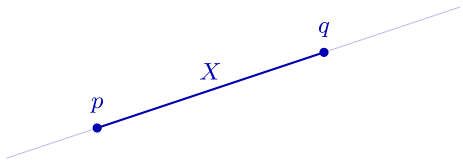


Figure 5: The relative interior of the closed segment $[p, q] \subset \mathbb{R}^2$ is the open segment $(p, q) = [p, q] \setminus \{p, q\}$, and its relative boundary is $\{p, q\}$. In contrast, $\text{int}([p, q]) = \emptyset$ and $\text{bd}([p, q]) = [p, q]$.

Definition 1.8. A point set X is called *affinely independent* if for any $x \in X$ we have $\text{aff}(X \setminus \{x\}) \neq \text{aff } X$. The points sets that are not affinely independent are called *affinely dependent* (Figure 6).

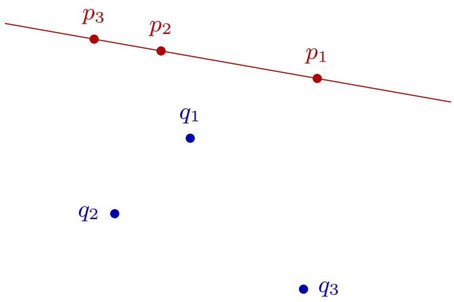


Figure 6: The set $\{p_1, p_2, p_3\}$ is affinely dependent, while $\{q_1, q_2, q_3\}$ is affinely indepentent.

Definition 1.9. Let $p_1, p_2, \ldots, p_k \in \mathbb{R}^n$ finitely many points, and let $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ be real numbers satisfying $\sum\limits_{i=1}^{k} \lambda_i = 1$. Then the point $\sum\limits_{i=1}^{k} \lambda_i p_i$ is called an *affine combination* of the points $p_1, p_2, \ldots, p_k$.

Proposition 1.10. *The affine hull of a set X is the set of the affine combinations of all finite point sets from X.*

Proof. Let Y denote the set of all affine combinations of finitely many points in X, and let $p \in X$ be an arbitrary point. Consider the points $p_1 = p, p_2, \ldots, p_k \in X$ and numbers $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ for which $\sum_{i=1}^{k} \lambda_i = 1$ is satisfied. According to our conditions:

$$\sum_{i=1}^{k} \lambda_i p_i = p_1 + \sum_{i=1}^{k} \lambda_i (p_i - p_1).$$

Thus the affine combination can be written as a translate of the point p with a linear combination of the vectors $p_i - p$. Hence, if L denotes the linear subspace formed by the linear combinations of the vectors $q - p$, $q \in X$, then $Y = p + L$. As it is clearly an affine subspace, we have $\text{aff}(X) \subseteq Y$.

On the other hand, if an affine subspace contains X, then it can be written in the form $p + L$ for some linear subspace L. The subspace L contains all vectors of the form

$q-p$, $q \in X$, and thus it contains their linear combinations as well. Hence, $p+L$ contains all affine combinations of points of X in the case that p is one of the points. Since any k-point affine combination is also a $(k+1)$-point affine combination in which one of the points is p, we have that $p+L$ contains all affine combinations of the points of X. Thus, $Y \subseteq p+L$, implying $Y \subseteq \operatorname{aff}(X)$. □

Corollary 1.11. *A point set X is affinely independent if and only if there is no point of X that can be written as an affine combination of some other points from X.*

Theorem 1.12. *Let $X = \{p_1, p_2, \ldots, p_k\} \subset \mathbb{R}^n$. Then X is affinely independent if and only if $\sum_{i=1}^{k} \lambda_i p_i = 0$ and $\sum_{i=1}^{k} \lambda_i = 0$ implies $\lambda_i = 0$ for all values of i.*

Proof. Assume that a point, say p_k, can be written as an affine combination of the other points; that is, $p_k = \sum_{i=1}^{k-1} \lambda_i p_i$, where $\sum_{i=1}^{k-1} \lambda_i = 1$. Then, setting $\lambda_k = -1$, we have $0 = \sum_{i=1}^{k} \lambda_i p_i$ and $\sum_{i=1}^{k} \lambda_i = 0$.

On the other hand, assume that for some values of the coefficients λ_i, not all of them zero, we have $0 = \sum_{i=1}^{k} \lambda_i p_i$ and $\sum_{i=1}^{k-1} \lambda_i = 0$. Without loss of generality, we may assume that $\lambda_k \neq 0$. For any $1 \leq i \leq k-1$, let $\lambda'_i = -\frac{\lambda_i}{\lambda_k}$. Then

$$\sum_{i=1}^{k-1} \lambda'_i = -\frac{\sum_{i=1}^{k-1} \lambda_i}{\lambda_k} = -\frac{-\lambda_k}{\lambda_k} = 1,$$

and

$$\sum_{i=1}^{k-1} \lambda'_i p_i = -\frac{1}{\lambda_k} \sum_{i=1}^{k-1} \lambda_i p_i = -\frac{1}{\lambda_k} (-\lambda_k p_k) = p_k,$$

and the point set is affinely dependent. □ □

Corollary 1.13. *If $X \subset \mathbb{R}^n$ is affinely independent, then every point of $\operatorname{aff}(X)$ can be uniquely written as an affine combination of some points in X.*

Theorem 1.14. *If $|X| \geq n+2$, then X is affinely dependent.*

Proof. Assume that $p_1, p_2, \ldots, p_{n+2} \in X$. Consider the vectors $p_2 - p_1, \ldots, p_{n+2} - p_1$ (Figure 7). Since the n-dimensional Euclidean space is an n-dimensional vector space, the above vectors are linearly dependent, that is one of them, say $p_{n+2} - p_1$, can be written as a linear combination of the other vectors: $p_{n+2} - p_1 = \sum_{i=2}^{n+1} \lambda_i (p_i - p_1)$. Let $\lambda_1 = 1 - \sum_{i=2}^{n+1} \lambda_i$. Then clearly $\sum_{i=1}^{n+1} \lambda_i = 1$. On the other hand,

$$p_{n+2} = p_1 + \sum_{i=2}^{n+1} \lambda_i (p_i - p_1) = \left(1 - \sum_{i=2}^{n+1} \lambda_i\right) p_1 + \sum_{i=2}^{n+1} \lambda_i p_i = \sum_{i=1}^{n+1} \lambda_i p_i,$$

that is, X is affinely dependent. □

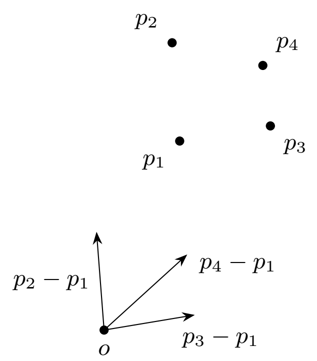


Figure 7: Any $n+2$ points $p_1, p_2, \ldots, p_{n+2}$ in $\mathbb{R}^n$ are affinely dependent, since the $n+1$ difference vectors $p_2 - 1, \ldots, p_{n+2} - p_1$ are necessarily linearly independent.

Corollary 1.15. *Every affine subspace of the space $\mathbb{R}^n$ is the affine hull of a most $n+1$ points.*

2 Convex combination, convex hull

We continue with a new topic.

Definition 2.1. Let $p_1, p_2, \ldots, p_k \in \mathbb{R}^n$. If a point p can be written in the form $\sum_{i=1}^{k} \lambda_i p_i$, $\sum_{i=1}^{k} \lambda_i = 1$, where $\lambda_i \geq 0$ for all is, then we say that p is a *convex combination* of the points $p_1, p_2, \ldots, p_k$.

Definition 2.2. The set of the convex combinations of the points $p, q \in \mathbb{R}^n$ is called the *closed segment* with endpoints p and q. If $p \neq q$, then the set $[p, q] \setminus \{p, q\}$ is called the *open segment* with endpoints p and q, and it is denoted by (p, q).

Definition 2.3. Let $K \subseteq \mathbb{R}^n$. The set K is called *convex*, if for arbitrary $p, q \in K$ we have $[p, q] \subseteq K$ (see Figures 8 and 9).

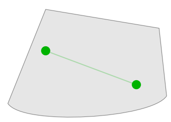


Figure 8: A convex set. The line segment joining any two of its points is also contained in the set.

Remark 2.4. *The intersection of arbitrarily many convex sets is convex (Figure 10).*

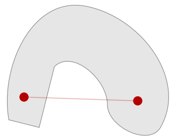


Figure 9: A set that is not convex. The line segment joining the two marked points is not contained in the set.

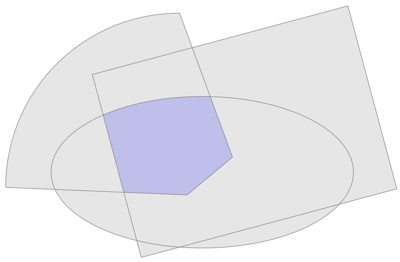


Figure 10: The intersection of arbitrarily many convex sets is convex.

Theorem 2.5. *Let $K \subseteq \mathbb{R}^n$ be a closed, convex set. Then K coincides with the intersection of the closed half spaces containing K.*

Proof. Let K' denote the intersection of the closed half spaces containing K. Clearly, $K \subseteq K'$. We need to show that $K' \subseteq K$.

Suppose for contradiction that there is some point $p \in K' \setminus K$. Consider the function $q \mapsto \text{dist}(p, q)$. We show that this function attains its minimum on K. If K is bounded, then it is compact, and thus the statement follows from the continuity of the distance function. If K is not bounded, then let us choose a closed ball B centered at p that contains a point from K. By the compactness of $K \cap B$ the function $\text{dist}(P, .)$ attains its minimum on $(K \cap B)$, and this minimum coincides with the minimum attained on K.

Let $\text{dist}(p, q)$ be the minimum of the function $\text{dist}(p, .)$, where $q \in K$, and let H denote the hyperplane containing q and perpendicular to $q - p$. Since the minimum is positive by the choice of p, $p \notin H$. On the other hand, if the open half space bounded by H and containing q contains some point $r \in K$, then the segment $[q, r]$, which belongs to K by the convexity of K, contains a point of K closer to p than q, which contradicts the choice of q (Figure 11). Thus, the closed half space bounded by H and not containing p contains K, which contradicts the choice of p.

□

It is easy to see that the closure of a convex set is convex. This yields the following remark.

Corollary 2.6. *If $K \subseteq \mathbb{R}^n$ convex, then for every boundary point of K there is a hyperplane H containing it such that K is contained in one of the two closed half spaces bounded by H.*

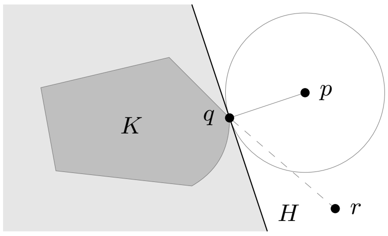


Figure 11: The set K is closed, therefore it has a point $q \in K$ closest to p. Since $p \notin K$, $q \neq p$. The hyperplane H through q and perpendicular to $q - p$ divides the space into two half spaces. If there was a point $r \in K$ in the open half space containing p then, by convexity, K would also have a point closer to p than q, a contradiction.

Definition 2.7. Let $X \subset \mathbb{R}^n$ be a nonempty set. Then the intersection of all convex sets that contain X is called the *convex hull* of X, and is denoted by $\text{conv}(X)$ (Figure 12).

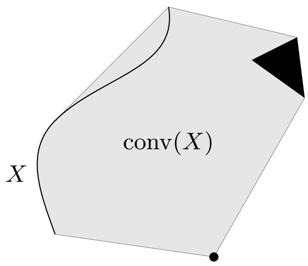


Figure 12: The shaded region is the convex hull of X (black).

Theorem 2.8. *Let $X \subset \mathbb{R}^n$ be a nonempty set. Then the convex hull of X is the set of the convex combinations of finite subsets of X.*

Proof. Let $p = \sum_{i=1}^{k} \lambda_i a_i$ and $q = \sum_{j=1}^{m} \mu_j b_j$ be two arbitrary convex combinations of points from X. Then a point of the segment $[p, q]$ can be written as $s = \alpha p + \beta q$ for some $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. But then $s = \alpha \sum_{i=1}^{k} \lambda_i a_i + \beta \sum_{j=1}^{m} \mu_j b_j = \sum_{i=1}^{k} \alpha\lambda_i a_i + \sum_{j=1}^{m} \beta\mu_j b_j$, which is a convex combination of the points $a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_m$, and hence, the set of convex combinations is convex.

Now, by induction on the number k of points, we prove that any convex set K containing X contains all convex combinations of points of X. Since points of a segment are convex combinations of the endpoints, for $k = 2$ the statement is follows from the convexity of X. Assume that K contains all k-element convex combinations, and consider some convex combination $p = \sum_{i=1}^{k+1} \alpha_i a_i$. If a coefficient in it is zero, we can apply the induction hypothesis directly. Thus, we may assume that e.g. $0 < \alpha_{k+1} < 1$. Then, let $\beta_i = \frac{\alpha_i}{1-\alpha_{k+1}}$ for all $i = 1, 2, \ldots, k$. Note that due to $\sum_{i=1}^{k} \beta_i = 1$, the point $q = \sum_{i=1}^{k} \beta_i a_i$ is an element of K. As $p = (1-\alpha_{k+1})q + \alpha_{k+1}a_{k+1}$ is a point of the segment $[q, a_{k+1}]$, we also have $p \in K$. □

3 Radon's, Carathéodory's and Helly's theorems

We continue the class with proving three fundamental theorems of convex geometry: Radon's, Carathéodory's and Helly's theorems.

Theorem 3.1 (Radon). *Let $X \subset \mathbb{R}^n$ be a set containing at least $n+2$ points. Then X can be decomposed into two parts whose convex hulls have a nonempty intersection.*

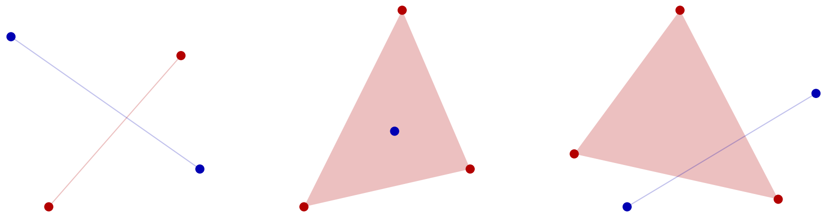


Figure 13: Illustration of Radon's theorem in the plane with 4 and 5 points.

Proof. Let $p_1, p_2, \ldots, p_m \in X$, where $m > n+1$. Consider the following homogeneous system of linear equations:

$$\sum_{i=1}^{m} \alpha_i = 0$$
$$\sum_{i=1}^{m} \alpha_i p_i = 0$$

This system of equations consists of $n+1$ equations and $m > n+1$ variables, and hence it has a nontrivial solution $(\beta_1, \beta_2, \ldots, \beta_m)$.

Let $V = \{i | \beta_i > 0\}$ and $W = \{i | \beta_i \leq 0\}$. Observe that because of the first equation of the system we have $V \neq \emptyset \neq W$, as otherwise $\beta_i = 0$ for all values of i, but the solution is nontrivial. We can also observe that by the same equation $\sum\limits_{i \in V} \beta_i = \sum\limits_{i \in W} (-\beta_i)$. Let $\beta > 0$ denote the common value of the two sides in the above equation. Then the point

$$p = \sum_{i \in V} \frac{\beta_i}{\beta} p_i = \sum_{i \in W} \frac{-\beta_i}{\beta} p_i$$

can be written as convex combinations of points from both $\{p_i | i \in V\}$,and $\{p_i | i \in W\}$, and thus, it lies in the intersection of the convex hulls of these two disjoint sets. □

It can be easily shown that if X is an affinely independent set of $n+1$ points for which $\operatorname{aff} X = \mathbb{R}^n$, then for X the above statement does not hold. Thus, the quantity $n+2$ in the theorem cannot be replaced by $n+1$.

Theorem 3.2 (Carathéodory)**.** *Let $X \subset \mathbb{R}^n$ be an arbitrary nonempty set. If $p \in \operatorname{conv} X$, then X has a subset Y consisting of at most $n+1$ points, satisfying $p \in \operatorname{conv}(Y)$.*

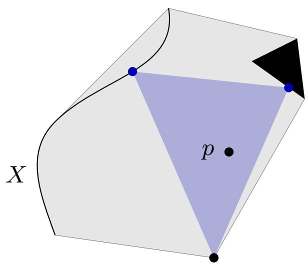


Figure 14: The point p in the convex hull of $X \subseteq \mathbb{R}^2$ can be expressed as a convex combination of three points of X.

Proof. Assume that $m > n+1$ is the smallest positive integer for which p can be written as a convex combination of m points of X. Let

$$p = \sum_{i=1}^{m} \alpha_i p_i, \tag{1}$$

where $\sum_{i=1}^{m} \alpha_i = 1$, and for $i = 1, 2, \ldots, m$ we have $\alpha_i \geq 0$ and $p_i \in X$. Since m is the smallest positive integer satisfying these conditions, we have $\alpha_i > 0$ for all values of i.

By Radon's theorem, the set $\{p_i | i = 1, 2, \ldots, m\}$ can be decomposed into two disjoint sets whose convex hulls have nonempty intersection. In other words, there are disjoint sets V and W for which $V \cup W = \{1, 2, \ldots, m\}$, and nonnegative numbers β_i for which $\sum_{i \in V} \beta_i = \sum_{i \in W} \beta_i = 1$ and $\sum_{i \in V} \beta_i p_i = \sum_{i \in W} \beta_i p_i$. Thus, by introducing the notation $\gamma_i = \beta_i$ for $i \in V$ and $\gamma_i = -\beta_i$ for $i \in W$, we obtain

$$\sum_{i=1}^{m} \gamma_i p_i = 0, \quad \text{and} \quad \sum_{i=1}^{m} \gamma_i = 0. \tag{2}$$

Let k be an index such that $\gamma_k < 0$ and

$$\frac{\alpha_k}{\gamma_k} \geq \frac{\alpha_i}{\gamma_i} \tag{3}$$

for all value of i with $\gamma_i < 0$.

Adding $\left(-\frac{\alpha_k}{\gamma_k}\right)$ times the equation (2) to (1), we obtain a linear combination

$$p = \sum_{i=1}^{m} \left(\alpha_i - \frac{\alpha_k}{\gamma_k}\gamma_i\right) p_i$$

in which the sum of the coefficients is 1. On the other hand, every coefficient is nonnegative, since it is clearly satisfied if $\gamma_i \geq 0$, and in the opposite case it is the consequence of the inequality in (3). As the kth coefficient is zero, we expressed p as a convex combination of at most $m-1$ points, which is a contradiction. □

Observe that if $X = \{p_1, p_2, \dots, p_{n+1}\}$ is affinely independent in $\mathbb{R}^n$, then the point $p = \frac{1}{n+1} \sum\limits_{i=1}^{n+1} p_i$ is in $\operatorname{conv}(X)$, but it is not contained in the convex hull of any proper subset of X. We can also observe that while Carathéodory's theorem describes how one can build up the convex hull of a set 'from inside', that is from the points of the set, Theorem 2.5 and Corollary 2.6 describe how to get to the convex hull 'from outside'.

Definition 3.3. The convex hulls of k-element subsets of $\mathbb{R}^n$ with $k \le n+1$ are called *simplices*. If the point set is affinely independent, we call the simplex *nondegenerate*. Then the elements of the point set are the *vertices* of the nondegenerate simplex, and the convex hull of two vertices is an *edge* of the simplex. If $k = n+1$, then the convex hull of n vertices is a *facet* of the simplex. If all edges of a nondegenerate simplex are of equal length, we call the simplex *regular*.

In the following we introduce an application of Carathéodory's theorem.

Theorem 3.4. *Let $H \subset \mathbb{R}^n$ be compact. Then $\operatorname{conv}(H)$ is also compact.*

Proof. Let

$$\Delta = \left\{ (\alpha_1, \alpha_2, \dots, \alpha_{n+1}) \in \mathbb{R}^{n+1} \,\middle|\, \sum_{i=1}^{n+1} \alpha_i = 1 \text{ and } \alpha_i \ge 0, i = 1, 2, \dots, n+1 \right\}.$$

Observe that Δ is compact. Consider the map $f : \mathbb{R}^{n+1} \times (\mathbb{R}^n)^{n+1} \to \mathbb{R}^n$ defined as

$$f(\alpha_1, \dots, \alpha_{n+1}, p_1, \dots, p_{n+1}) = \sum_{i=1}^{n+1} \alpha_i p_i$$

for all $\alpha_i \in \mathbb{R}, p_i \in \mathbb{R}^n$ $(i = 1, 2, \dots, n+1)$.

Then f is a continuous map and $f(\Delta \times H^{n+1}) = \operatorname{conv} H$. As the direct product of compact sets is compact, the image of a compact set under a continuous map is compact, we have that $\operatorname{conv}(H)$ is compact. □

Before describing another application of Carathéodory's theorem, we verify another statement that can often be used in convex geometry problems.

Proposition 3.5. *Let H be a closed half space bounded by the hyperplane H_0, and let $X \subset H$ be arbitrary. Then $\operatorname{conv}(X) \cap H_0 = \operatorname{conv}(X \cap H_0)$ (Figure 15).*

Proof. Since H_0 is convex and $X \cap H_0 \subseteq X$, we obtain $\operatorname{conv}(X \cap H_0) \subseteq \operatorname{conv}(X) \cap H_0$. We show that $\operatorname{conv}(X) \cap H_0 \subseteq \operatorname{conv}(X \cap H_0)$.

Let $p \in \operatorname{conv}(X) \cap H_0$ be arbitrary. Then, by Theorem 2.8 with a suitable choice of $p_1, \dots, p_k \in X$, $\alpha_1, \dots, \alpha_k > 0$, $\sum_{i=1}^k \alpha_i = 1$, we have $p = \sum_{i=1}^k \alpha_i p_i$. As H is a closed

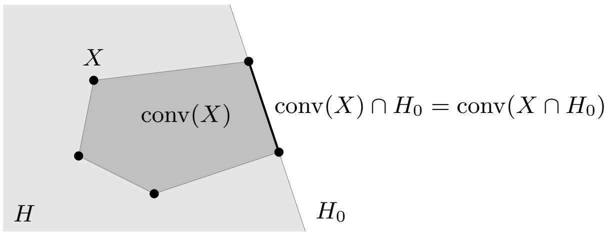


Figure 15: If X is contained in one of the half spaces bounded by the hyperplane H_0, then the intersection of its convex hull with H_0 is $\mathrm{conv}(X \cap H_0)$

half space, there are some $\alpha \in \mathbb{R}$ and $u \in \mathbb{R}^n$ such that $H = \{x \in \mathbb{R}^n | \langle x, u \rangle \geq \alpha\}$ and $H_0 = \{x \in \mathbb{R}^n | \langle x, u \rangle = \alpha\}$. Thus,

$$\alpha = \langle u, p \rangle = \left\langle u, \sum_{i=1}^{k} \alpha_i p_i \right\rangle = \sum_{i=1}^{k} \alpha_i \langle u, p_i \rangle \geq \sum_{i=1}^{k} \alpha_i \alpha = 1,$$

with equality if and only if $\langle u, p_i \rangle = \alpha$ for all values of i. Consequently, $p_i \in H_0 \cap X$ for all is, from which $p \in \mathrm{conv}(X \cap H_0)$. □

Theorem 3.6 (colorful Carathéodory theorem)**.** *Let $X_1, X_2, \ldots, X_{n+1} \subset \mathbb{R}^n$ be compact sets. Assume that for any i we have $o \in \mathrm{conv}\, X_i$. Then there are some points $p_i \in X_i$ such that $o \in \mathrm{conv}\{p_1, p_2, \ldots, p_{n+1}\}$ (see Figures 16 and 17).*

In the theorem, X_i denotes the set of points with 'color i'. Thus, the statement guarantees that there is a 'rainbow simplex' containing the origin.

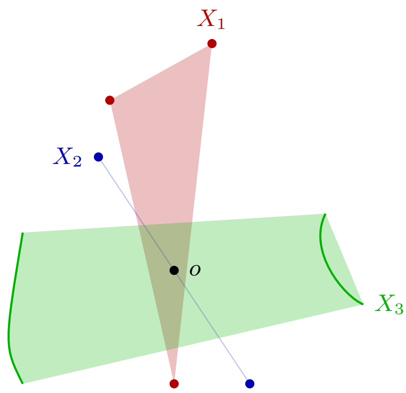


Figure 16: The origin is contained in $\mathrm{conv}(X_i)$ for all i.

Proof. We prove by contradiction. Suppose that there is no 'rainbow simplex' containing the origin. Let $Y = \mathrm{conv}(p_1, p_2, \ldots, p_{n+1})$, $p_i \in X_i$ be a 'rainbow simplex' whose distance from o is minimal. Since the sets X_i are compact, such a simplex exists. Let q be the (unique) point of Y whose distance from o is minimal, and let H denote the closed half

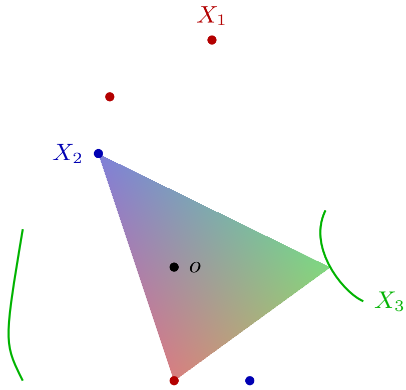


Figure 17: A rainbow simplex containing the origin.

space not containing o, which contains q in its boundary and whose bounding hyperplane is perpendicular to q. If Y had a point in the complement of H, then Y would contain a point closer to o than q, and thus, $Y \subset H$ (Figure 18).

If Y had a vertex p_i which is not in the boundary of H, then $o \in \operatorname{conv} X_i$ yields that there is some point $p_i' \in X_i$ not in H. But then $q \in \operatorname{conv}\{p_1, p_2, \ldots, p_{i-1}, p_{i+1}, \ldots, p_{n+1}\}$ by Proposition 3.5, and hence $\operatorname{conv}(p_1, \ldots, p_{i-1}, p_i', p_{i+1}, \ldots, p_{n+1})$ is a simplex which has a point closer to o than q, a contradiction. Thus Y is contained in the bounding hyperplane of H. But then, applying Carathéodory's theorem for this hyperplane, we obtain that for a suitable index i, we have that $q \in \operatorname{conv}\{p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_{n+1}\}$, and thus, similar to the previous case, we may replace p_i with a point $p_i' \in X_i$ in the complement of H, we obtain a simplex closer to o. □

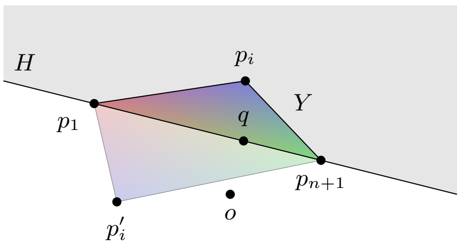


Figure 18: If Y is a rainbow simplex not containing the origin, then it is contained in the half space H. In the proof of Theorem 3.6 we show that in this case there is a vertex p_i that may be replaced with $p_i' \notin H$, leading to a rainbow simplex closer to o.

We continue with the description of an important theorem of convex geometry, and with an introduction of one of its applications.

Theorem 3.7 (Helly, finite)**.** *Let $\mathcal{K}$ be a finite family of at least $n+1$ convex sets in $\mathbb{R}^n$. If any $(n+1)$ elements of $\mathcal{K}$ have a nonempty intersection, then all elements of $\mathcal{K}$ have a nonempty intersection.*

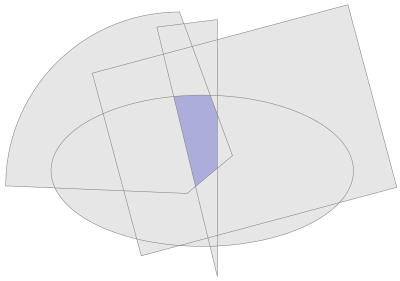


Figure 19: Helly's theorem in the plane. If any three members of a finite family of convex sets have a nonempty intersection, then the family itself has nonempty intersection.

Proof. Let the cardinality of $\mathcal{K}$ be $|\mathcal{K}| = k$. We prove the theorem by induction on k. The statement clearly holds if $k = n + 1$. Let us assume that it holds for all families with k elements, and let us consider a family $\mathcal{K}$ consisting of $k + 1$ convex sets in $\mathbb{R}^n$ with the property that any $n + 1$ elements of $\mathcal{K}$ have a nonempty intersection. By the induction hypothesis, for any $K \in \mathcal{K}$ there is a point p_K with the property that p_K is contained in every element of $\mathcal{K}$ but K. Let $X = \{p_K | K \in \mathcal{K}\}$ (Figure 20).

Radon's theorem implies that X can be written as the disjoint union of two sets X_1, X_2, whose convex hulls have a nonempty intersection. Let $p \in \operatorname{conv} X_1 \cap \operatorname{conv} X_2$. As $p_K \in K'$ for every $K' \neq K$, $K' \in \mathcal{K}$, we have that if $p_K \in X_1$, then $X_2 \subset K$. This yields by the convexity of K that $\operatorname{conv} X_2 \subset K$. We obtain similarly that if $p_K \in X_2$, then $\operatorname{conv} X_1 \subset K$. Now, since $p \in \operatorname{conv} X_1 \cap \operatorname{conv} X_2$, from this it follows that $p \in K$ for every $K \in \mathcal{K}$; that is, the intersection of all elements of $\mathcal{K}$ is not empty. □

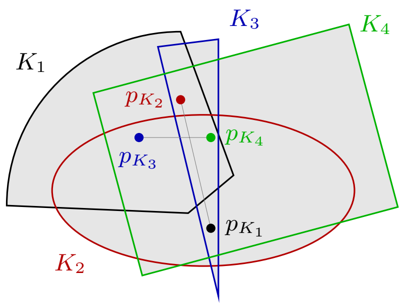


Figure 20: Illustration of the induction step in the proof of Theorem 3.7. Any three of the convex sets K_1, K_2, K_3, K_4 intersect, therefore there exist points p_{K_i} contained in all except possibly in K_i. By Radon's theorem, the set of these four points can be partitioned into two subsets with nonempty intersection. Any point in the intersection is contained in $K_1 \cap K_2 \cap K_3 \cap K_4$.

The example of the $n+1$ facets of a simplex shows that there are families of convex sets in $\mathbb{R}^n$ in which every n elements have a nonempty intersection, but there is no point contained in all elements of the family.

We have seen that Radon's theorem implies both Carathéodory's and Helly's theorem. Nevertheless, it can be shown that the Radon's theorem can be derived from any of the two latter theorems, which implies that these theorems are equivalent.

Helly's theorem also has a variant for families with infinitely many members.

Theorem 3.8 (Helly, infinite)**.** *Let $\mathcal{K}$ be a family of at least $n+1$ closed, convex sets in $\mathbb{R}^n$ such that at least one member of $\mathcal{K}$ is compact. Assume that any $n+1$ elements of $\mathcal{K}$ have a nonempty intersection. Then there is a point which is contained in every element of $\mathcal{K}$.*

Proof. According to the previous theorem it is sufficient to examine families $\mathcal{K}$ with infinitely many members, and we can also assume that any finitely many elements of $\mathcal{K}$ have a nonempty intersection. Assume that there is no point belonging to every element of $\mathcal{K}$. Let $K \in \mathcal{K}$ be a compact, closed set. Observe that all elements of the family $\mathcal{K}' = \{\mathbb{R}^n \setminus C | C \in \mathcal{K}\}$ are open. On the other hand, since there is no point that belongs to every member of $\mathcal{K}$, the family $\mathcal{K}'$ is an open cover of $\mathbb{R}^n$, and in particular, K. As K is compact, $\mathcal{K}'$ has finitely many elements whose union covers K. But then the complements of these sets has no common point that belongs to K, which contradicts our assumption that any finitely many elements of $\mathcal{K}$ have a common point. □

Our next examples show that the statement in the theorem does not hold if $\mathcal{K}$ has elements that are not closed, or if $\mathcal{K}$ has no compact element.

Example 3.9.

(i) *Let $K_i = \left\{(x,y) \in \mathbb{R}^2 \middle| \left(x - \frac{1}{i}\right)^2 + y^2 \leq \frac{1}{i^2}\right\}$ for every $i = 1,2,3,\ldots$, and let $K_0 = \left\{(x,y) \in \mathbb{R}^2 \middle| (x-2)^2 + y^2 < 4\right\}$ (Figure 21). It can be easily seen that any finitely many elements of the family $\mathcal{K} = \{K_i | i = 0,1,2,\ldots\}$ have a nonempty intersection, but the intersection of all elements is the empty set.*

(ii) *Let $K_i = \left\{(x,y) \in \mathbb{R}^2 \middle| y \geq i\right\}$ be for every $i = 1,2,3,\ldots$ (Figure 22). Then any finitely many elements of $\mathcal{K} = \{K_i | i = 1,2,\ldots\}$ have a nonempty intersection, but the intersection of all elements is empty.*

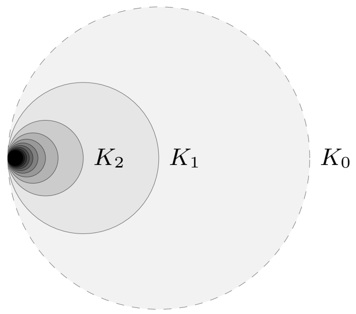


Figure 21: An infinite family of bounded convex sets with empty intersection: K_i is the closed disk with center $(\frac{1}{i}, 0)$ and radius $\frac{1}{i}$, K_0 is the open disk with center $(2,0)$ and radius 2.

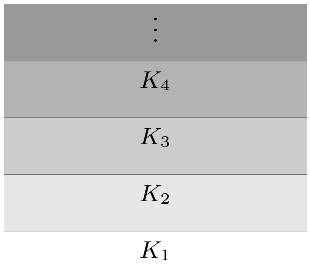


Figure 22: An infinite family $(K_i)_{i=1}^{\infty}$ of closed half planes with the property that any finite subfamily has nonempty intersection, but $\bigcap_{i=1}^{\infty} K_i = \emptyset$.

4 Jung's theorem and Minkowski sums

We present an application of Helly's theorems.

Definition 4.1. The *diameter* of a bounded set $A \subset \mathbb{R}^n$ is the supremum of the distances of all pairs of points from the set.

Theorem 4.2 (Jung)**.** *A set in $\mathbb{R}^n$ having diameter d can be covered by a closed Euclidean ball of radius $d \cdot \sqrt{\frac{n}{2(n+1)}}$.*

We remark that the quantity in the theorem is the circumradius of the regular n-dimensional simplex of edge length d.

Proof. Let the diameter of $S \subset \mathbb{R}^n$ be d, and for every $p \in S$, let G_p denote the set of points x in $\mathbb{R}^n$ with the property that the closed ball of radius $d \cdot \sqrt{\frac{n}{2(n+1)}}$ and center x covers p. Note that G_p is the closed ball of radius $d \cdot \sqrt{\frac{n}{2(n+1)}}$ centered at p (both conditions are equivalent to saying that $\|x - p\| \leq d \cdot \sqrt{\frac{n}{2(n+1)}}$), and thus, it is compact and convex (Figure 23). Hence, if we can verify that $\bigcap_{i=1}^{k} G_{p_i} \neq \emptyset$ for any $p_1, p_2, \ldots, p_k \in S$ and $k \leq n+1$, then from Helly's theorem (infinite version) it follows that $\bigcap_{p \in S} G_p \neq \emptyset$, which readily yields our theorem.

Let $p_1, p_2, \ldots, p_k \in S$ with $k \leq n+1$, and let q be the center of the smallest closed ball containing the points $p_1, p_2, \ldots, p_k$. We show that the radius of G is at most $r \leq d\sqrt{\frac{n}{2(n+1)}}$. Observe that $q \in \operatorname{conv}\{p_1, p_2, \ldots, p_k\}$, as otherwise there is a smaller ball that contains the points (Figure 24). In addition, since we have only finitely many points, G is the smallest ball that contains those p_is that are contained in its boundary, and thus, we may assume that $\|p_i - q\| = r$ for all values of i. Let $v_i = p_i - q$ for $i = 1, \ldots, k$. Then $o \in \operatorname{conv}\{v_1, v_2, \ldots, v_k\}$ As the diameter of S is d, we have $\|v_i - v_j\| = \operatorname{dist}(p_i, p_j) \leq d$ for all i and j (Figure 25). Write $o = \sum_{i=1}^{k} \alpha_i v_i$, where $\alpha_i \geq 0$

and $\sum_{i=1}^{k} \alpha_i = 1$. For all i we have

$$\begin{aligned}\alpha_i r^2 = \alpha_i \|v_i\|^2 = \langle \alpha_i v_i, v_i \rangle = \langle -\sum_{j\neq i} \alpha_j v_j, v_i \rangle = -\sum_{j\neq i} \alpha_j \langle v_j, v_i \rangle \\ = -\sum_{j\neq i} \alpha_j \frac{1}{2}\left(\|v_i\|^2 + \|v_j\|^2 - \|v_i - v_j\|^2\right) = \sum_{j\neq i} \alpha_j \left(\frac{1}{2}\|v_i - v_j\|^2 - r^2\right) \\ \leq \sum_{j\neq i} \alpha_j \left(\frac{d^2}{2} - r^2\right) = (1-\alpha_i)\left(\frac{d^2}{2} - r^2\right),\end{aligned}$$

therefore

$$r^2 \leq (1-\alpha_i)\frac{d^2}{2}.$$

Choosing an index i such that $\alpha_i \geq \frac{1}{k}$, we obtain

$$r^2 \leq \frac{k-1}{2k} d^2,$$

from which, as $k \leq n+1$, the inequality

$$r^2 \leq \frac{k-1}{2k} d^2 \leq \frac{n}{2n+2} d^2$$

follows. □

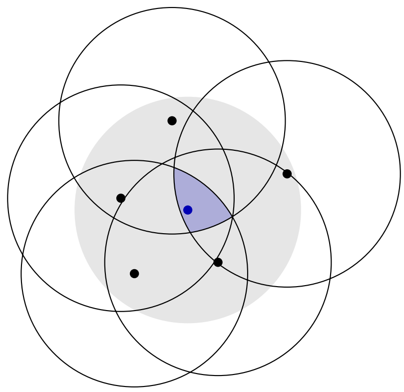


Figure 23: A set of points can be covered with a closed ball of radius R (shaded area) if and only if the family of balls of radius R, centered at the points, has a nonempty intersection.

5 Minkowski sum and support function

To continue, recall the definition of the Minkowski sum of two sets from the first lecture.

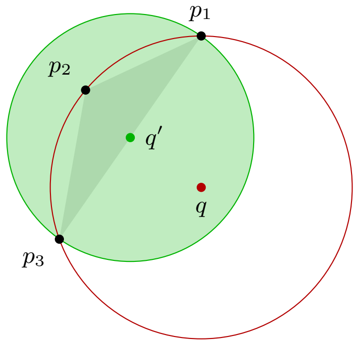


Figure 24: The point q is not contained in the convex hull of $\{p_1, p_2, p_3\}$, therefore it is not the center of the smallest ball containing these points. The green ball (centered at q') has smaller radius and contains all three points.

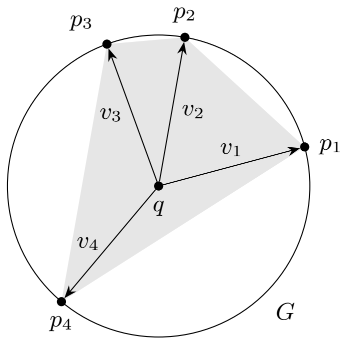


Figure 25: G is the smallest ball containing the points p_i on its boundary. The vectors $v_i = p_i - q$ satisfy $\|v_i\| = r$, $\|v_i - v_j\| = \mathrm{dist}(p_i, p_j)$, and $o \in \mathrm{conv}\{v_1, \ldots, v_k\}$.

Definition 5.1 (Definition 1.1, repeated)**.** Let V_1 and V_2 be two point sets, and let $\lambda \in \mathbb{R}$. Then

$$V_1 + V_2 = \{v_1 + v_2 | v_1 \in V_1, v_2 \in V_2\}$$

is the Minkowski sum of the two sets and

$$\lambda V_1 = \{\lambda v_1 | v_1 \in V_1\}$$

is the multiple of V_1 by λ.

Remark 5.2. *To 'draw' the Minkowski sum of two sets we should think it over that by definition, $V_1 + V_2 = \bigcup_{v_1 \in V_1} (v_1 + V_2)$, implying that the sum of the two sets can be obtained as the region 'swept' by the translates of one of the sets where the translation vectors run over the other set (Figure 26).*

Proposition 5.3. *If $K, L \subset \mathbb{R}^n$ are convex, then $K + L$ is convex (Figure 27).*

Proof. We need to show that the segment connecting any two points of $K + L$ belongs to $K + L$. In other words, we need to show that if $p_K, q_K \in K$ and $p_L, q_L \in L$, then

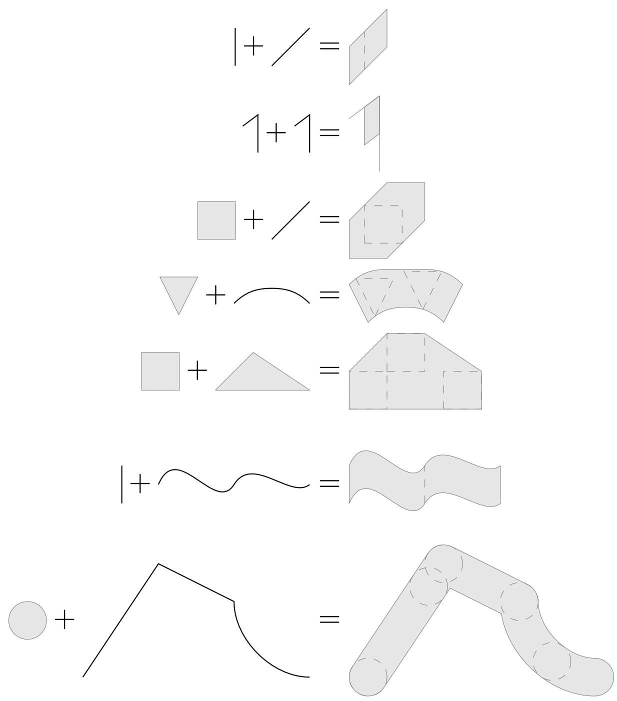


Figure 26: Examples of Minkowski sums

$[p_K+p_L, q_K+q_L] \subseteq K+L$. Let $t \in [0,1]$ be arbitrary. Then $t(p_K+q_K)+(1-t)(p_L+q_L) = (tp_K + (1-t)q_K) + (tp_L + (1-t)q_L)$, where by the convexity of K and L, we have $tp_K+(1-t)q_K \in K$ and $tp_L+(1-t)q_L \in L$. Thus, $t(p_K+q_K)+(1-t)(p_L+q_L) \in K+L$, from which the statement follows. $\square$

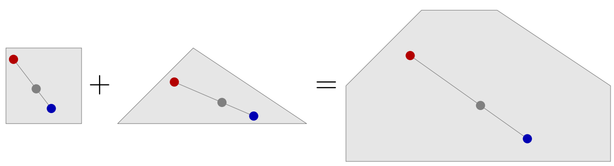


Figure 27: The Minkowski sum of two convex sets is convex

Definition 5.4. Let $A \subset \mathbb{R}^n$ be an arbitrary bounded set. Then the function

$$h_A : \mathbb{R}^n \to \mathbb{R}, \quad h_A(x) = \sup\{\langle x, y\rangle | y \in A\}$$

is called the *support function* of A.

Example 5.5.

(i) If $A \subseteq \mathbb{R}$ is bounded and $a = \inf A$, $b = \sup A$, then

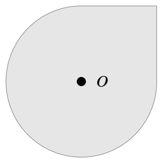
$$h_A(x) = \max\{ax, bx\}.$$

(ii) Let B denote the unit ball in $\mathbb{R}^n$. Then

$$h_B(x) = \sup\left\{\langle x, y\rangle | y \in \mathbb{R}^n, \|y\| \leq 1\right\} = \|x\|.$$

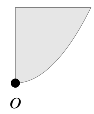
(iii) Let $A = \left\{(y_1, y_2) \in \mathbb{R}^2 \middle| y_1^2 + y_2^2 \leq 1\right\} \cup ([0,1] \times [0,1])$. Then

$$h_A(x) = \max\{\|x\|, x_1 + x_2\}.$$



(iv) Let $A = \left\{(y_1, y_2) \in \mathbb{R}^2 \middle| 0 \leq y_1^2 \leq y_2 \leq 1, 0 \leq y_1\right\} \subseteq \mathbb{R}^2$. Then

$$h_A(x) = \begin{cases} -\frac{x_1^2}{4x_2} & \text{if } 0 < x_1 \leq -2x_2 \\ \max\{0, x_2, x_1 + x_2\} & \text{otherwise} \end{cases}$$



Proposition 5.6. *Let $A, B \subseteq \mathbb{R}^n$ be bounded sets.*

(i) if $A \subseteq B$ then $h_A(x) \leq h_B(x)$ holds for all $x \in \mathbb{R}^n$,

(ii) if B is also closed and convex, and $h_A(x) \leq h_B(x)$ holds for all $x \in \mathbb{R}^n$, then $A \subseteq B$.

In particular, if A and B are compact and convex, then $A \subseteq B \iff \forall x \in \mathbb{R}^n : h_A(x) \leq h_B(x)$.

Proof. (i): Suppose that $A \subseteq B$. Then

$$\begin{aligned} h_A(x) &= \sup\left\{\langle x, y\rangle | y \in A\right\} \\ &\leq \sup\left\{\langle x, y\rangle | y \in B\right\} \\ &= h_B(x), \end{aligned}$$

since the supremum over a subset is at most the supremum over the larger set.

(ii): Suppose that B is closed and convex, and $A \not\subseteq B$. Then there is a point $p \in A \setminus B$. But then by Theorem 2.5 and Remark 1.5 there is some $u \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $\langle u, p\rangle > \alpha$, and $\langle u, x\rangle \leq \alpha$ for every $x \in B$ (see Figure 28). But from this $h_A(u) > h_B(u)$ follows. $\square$

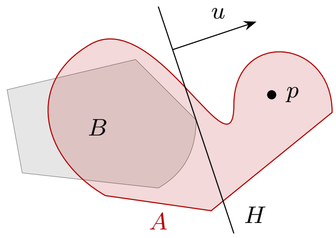


Figure 28: If B is closed and convex, and $p \in A \setminus B$, then there exist $u \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $\langle u, p \rangle > \alpha$, and $\langle u, x \rangle \leq \alpha$ for every $x \in B$. This implies $h_B(u) \leq \alpha < h_A(u)$.

Theorem 5.7. *Let $A \subset \mathbb{R}^n$ be an arbitrary bounded set containing o. Then the support function h_A of A is:*

(i) convex, that is, $h(tx + (1-t)y) \leq th(x) + (1-t)h(y)$ for every $x, y \in \mathbb{R}^n$ and $t \in [0, 1]$;

(ii) h nonnegative, and for any $\lambda \geq 0$ and $x \in \mathbb{R}^n$, we have $h(\lambda x) = \lambda h(x)$.

Furthermore, for any function h satisfying properties (i) and (ii), there is a unique compact, convex set $A \subset \mathbb{R}^n$, containing o, whose support function is h.

Proof. Clearly,

$$\begin{aligned} h_A(tx + (1-t)y) &= \sup\{\langle tx + (1-t)y, z \rangle | z \in A\} \\ &\leq t \sup\{\langle x, z \rangle | z \in A\} + (1-t) \sup\{\langle y, z \rangle | z \in A\} \\ &= th_A(x) + (1-t)h_A(y), \end{aligned}$$

that is, h_A is convex. The second property readily follows from the properties of the inner product.

Now, let h be a function satisfying properties (i) and (ii), and let

$$A = \{y \in \mathbb{R}^n | \langle x, y \rangle \leq h(x) \text{ for every } x \in \mathbb{R}^n\}.$$

As for any fixed x, the set of points y satisfying the inequality $\langle x, y \rangle \leq h(x)$ is a closed half space containing o, the set A, which is the intersection of such sets, is a closed, convex set containing o. Denoting by e_i the unit vector pointing in the ith coordinate direction, for all $y \in A$ the ith coordinate of y satisfies

$$\langle e_i, y \rangle \leq h(e_i)$$

and

$$\langle e_i, y \rangle = -\langle -e_i, y \rangle \geq -h(-e_i),$$

therefore A is bounded. Thus, we have seen that A is compact. On the other hand, for any vector $z \in \mathbb{R}^n$, we have $h_A(z) = \sup\{\langle z, y \rangle | y \in A\} \leq h(z)$ by the definition of A.

We will show that $h_A(z) \geq h(z)$, that is, that there is a point $y \in A$, for which $\langle y, z \rangle = h(z)$. Since this statement clearly holds if $z = o$ or $h(z) = 0$, we assume that $z \neq o$ and $h(z) > 0$. Let us define the *epigraph* of h as the closed set $E_h = \{(x, \alpha) | h(x) \leq \alpha\} \subseteq \mathbb{R}^n \times \mathbb{R}$ (note that this set is the region 'above' the graph of h in $\mathbb{R}^{n+1}$, see Figure 29). If $(x, \alpha), (y, \beta) \in E_h$ and $t \in [0, 1]$, then $h(tx + (1-t)y) \leq th(x) + (1-t)h(y) \leq t\alpha + (1-t)\beta$, implying that E_h is convex, and clearly, if $(x, \alpha) \in E_h$ and $\lambda \geq 0$, then $(\lambda x, \lambda\alpha) \in E_h$. By the definition of epigraph, $(z, h(z))$ is a boundary point of E_h, and hence, by Corollary 2.6 and Remark 1.5, there are $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$ and $\alpha \in \mathbb{R}$ which satisfy $\langle y, w \rangle + \beta\gamma \leq \alpha$ for any $(w, \gamma) \in E_h$, and $\langle y, z \rangle + \beta h(z) = \alpha$. Since $z \neq o$, from the positive homogeneity of E_h it follows that $\alpha = 0$. On the other hand, since h is defined on the whole space $\mathbb{R}^n$, we have $\beta \neq 0$, and thus, with a suitable choice of y we may assume that $\beta = -1$. But from this $\langle y, z \rangle = h(z)$, which is what we wanted to prove. Thus, $h_A = h$.

Finally, the support functions of different compact, convex sets containing o are different by Proposition 5.6. □

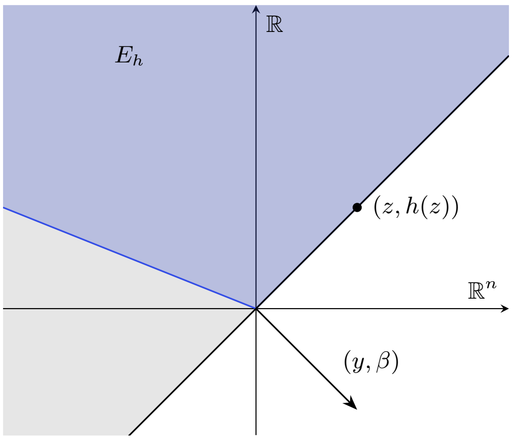


Figure 29: The epigraph of h (blue region) is contained in the closed half space (gray) with outer normal vector $(y, \beta) \in \mathbb{R}^n \times \mathbb{R}$.

Proposition 5.8. *For any bounded sets $K, L \subset \mathbb{R}^n$, we have $h_{K+L} = h_K + h_L$.*

Proof. If $x \in \mathbb{R}^n$, then

$$\begin{aligned} h_{K+L}(x) &= \sup\{\langle x, y \rangle + \langle x, z \rangle | y \in K, z \in L\} \\ &= \sup\{\langle x, y \rangle | y \in K\} + \sup\{\langle x, z \rangle | z \in L\} \\ &= h_K(x) + h_L(x). \end{aligned}$$

□

6 Separation

Remark 6.1. *Let $L_1, L_2 \subseteq \mathbb{R}^n$ be linear subspaces with* $\dim(L_1) = k$ *and* $\dim(L_2) = n-k$ *for some $0 \leq k \leq n$, and let $L_1 \cap L_2 = \{o\}$. Then the union of a basis of L_1 and a basis of L_2 is a basis of $\mathbb{R}^n$, and hence, for any point $p \in \mathbb{R}^n$ there are unique points $p_1 \in L_1$, $p_2 \in L_2$ satisfying $p = p_1 + p_2$.*

Definition 6.2. Let $L_1, L_2 \subseteq \mathbb{R}^n$ be linear subspaces with $\dim(L_1) = k$ and $\dim(L_2) = n - k$ for some $0 \leq k \leq n$, and let $L_1 \cap L_2 = \{o\}$. For any $x \in \mathbb{R}^n$ let $x_1 \in L_1$, $x_2 \in L_2$ denote those unique points that satisfy $x = x_1 + x_2$ (Figure 30). Then the linear transformation $\pi : \mathbb{R}^n \to L_2$, $\pi(x) = x_2$ is called *projection onto L_2 parallel to L_1*. If L_1 is the orthogonal complement of L_2, then we say that π is the *orthogonal projection onto L_2*.

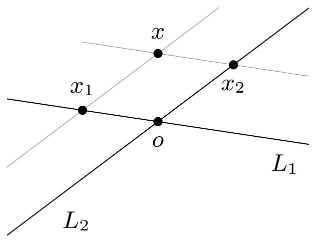


Figure 30: If the linear subspaces $L_1, L_2 \subseteq \mathbb{R}^n$ satisfy $L_1 \cap L_2 = \{0\}$ and $\dim(L_1) + \dim(L_2) = n$, then any vector $x \in \mathbb{R}^n$ can be uniquely decomposed as $x = x_1 + x_2$ where $x_1 \in L_1$ and $x_2 \in L_2$.

From the definition it is clear that if $\dim(L_1) = k$ and L is an affine subspace of dimension m in L_2, then $\pi^{-1}(L)$ is an $(m + k)$-dimensional affine subspace in $\mathbb{R}^n$.

Remark 6.3. *If the conditions of the previous remark are satisfied for the linear subspaces $L_1, L_2 \subseteq \mathbb{R}^n$ then for any $p_1, p_2 \in \mathbb{R}^n$, the intersection of $p_1 + L_1$ and $p_2 + L_2$ is a singleton. Indeed, by the previous remark, p_1 can be decomposed to the sum of a vector from L_1 and a vector from L_2,and hence, as $x + L_1 = L_1$ if $x \in L_1$, we may assume that $p_1 \in L_2$. Similarly, we may assume that $p_2 \in L_1$. Thus, if $x \in \mathbb{R}^n$ is contained in both subspaces, then, writing it in the form $x = x_1 + x_2$, $x_1 \in L_1$, $x_2 \in L_2$, the previous remark implies that $x_1 = p_2$ and $x_2 = p_1$; on the other hand $p_1 + p_2$ is an element of both subspaces. Based on this observation, projection can be defined not only for linear subspaces, but also for affine subspaces.*

Proposition 6.4. *Let $L_1, L_2 \subseteq \mathbb{R}^n$ be linear subspaces with* $\dim(L_1) = k$ *and* $\dim(L_2) = n-k$ *for some $0 \leq k \leq n$, and let $L_1 \cap L_2 = \{o\}$. Let π be the projection onto L_2 parallel to L_1. Then for any open/compact/convex set $X \subset \mathbb{R}^n$, $\pi(X)$ is open/compact/convex, respectively, and for any open/closed/convex set $Y \subseteq L_2$, the set $\pi^{-1}(Y)$ is open/closed/ convex, respectively.*

Proof. For any point $x \in \mathbb{R}^n$ the projection of a neighborhood of x is a neighborhood of $\pi(x)$ in L_2, and hence, if $X \subseteq \mathbb{R}^n$ open, then $\pi(X)$ is also open. Similarly, the projection of a closed segment is the closed segment connecting the projections of the endpoints, which yields that if X is convex, then so is $\pi(X)$. The statement for the projection of a compact set follows from the observation that π is a continuous function, and thus, the image of a compact set is compact. Similarly, it also follows that the preimage of an open/closed set is open/closed, respectively. Now, if $Y \subset L_2$ is convex, then for any $p, q \in Y$ choose some points $p', q' \in \mathbb{R}^n$ $\pi(p') = p$, $\pi(q') = q$. As $\pi([p', q']) = [p, q] \subseteq Y$ by the convexity of Y, we clearly have $[p', q'] \subseteq \pi^{-1}([p, q])$, implying that $\pi^{-1}(Y)$ is convex. □

Definition 6.5. Let $A, B \subseteq \mathbb{R}^n$. Let H be a hyperplane, and let H^+ and H^- be the two closed half spaces bounded by H. We say that H *separates* A and B if $A \subseteq H^+$ and $B \subseteq H^-$, or $B \subseteq H^+$ and $A \subseteq H^-$ (Figure 31). If H separates A and B, and $A \cap H = B \cap H = \emptyset$, then we say that H *strictly separates* A and B (Figure 32). If $A \subseteq H$, and $B \subseteq H^+$ or $B \subseteq H^-$, then we say that H *isolates* A from B (Figure 33). If, in addition, $B \cap H = \emptyset$, then we say that H *strictly isolates* A from B (Figure 34).

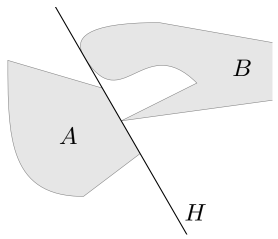


Figure 31: The hyperplane H separates A and B.

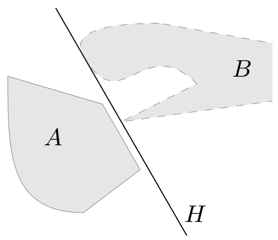


Figure 32: The hyperplane H strictly separates A and B.

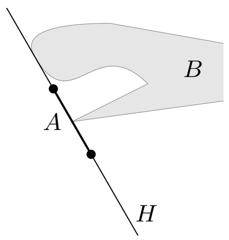


Figure 33: The hyperplane H isolates A from B.

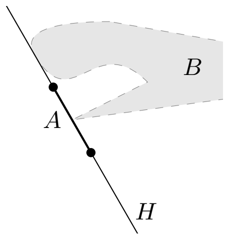


Figure 34: The hyperplane H strictly isolates A from B.

Theorem 6.6 (Isolation theorem)**.** *Let $K \subseteq \mathbb{R}^n$ be an open, convex set, and let $o \notin K$. Then there is a hyperplane H that isolates o from K.*

We remark that if a hyperplane H isolates of from K, then by the openness of K it also strictly isolates o from K.

Proof. The statement is trivial if $n = 1$. First, we prove it for $n = 2$. Let $\mathbb{S}^1$ be the set of unit vectors in $\mathbb{R}^2$, i.e. let it be the boundary of the circular disk centered at o and with unit radius. Let $p : \mathbb{R}^2 \setminus \{o\} \to \mathbb{S}^1$ be the central projection onto $\mathbb{S}^1$, i.e. let $p(v) = \frac{v}{||v||}$. Since K is convex, therefore it is connected, and thus, $p(K)$ is also connected (Figure 35). It is also clear that since K is open, the set $p(K)$ is also open. Thus, $p(K)$ is an open circular arc in $\mathbb{S}^1$. If $p(K)$ contains two opposite points $u, -u$, then there would be positive real numbers $\lambda_1, \lambda_2 > 0$ with $\lambda_1 u, -\lambda_2 u \in K$. But this would imply by the convexity of K that $o \in K$, which contradicts our assumptions. Hence, $p(K)$ does not contain opposite points, which yields that the length of $p(K)$ is at most π, or in other words, there are opposite points $u, -u \in \mathbb{S}^1$ such that neither one belongs to $p(K)$. This yields that there is a line through o disjoint from K.

If $n > 2$, we prove the statement by induction on n. Assume that the statement holds in $\mathbb{R}^k$ for every $1 \leq k < n$.

Consider a plane P through o. Since $P \cap K$ is an open, convex set, we may apply the case $n = 2$ of the statement and obtain a line $L \subset P$ through o disjoint from K. Let $H = L^\perp$ be the orthogonal complement of L. Let π be the orthogonal projection onto H parallel to L (Figure 36). Then by Proposition 6.4, $\pi(K)$ is an open, convex set in H, and thus, by the induction hypothesis, there is some $(n-2)$-dimensional linear subspace $G \subset H$ $(n-2)$ disjoint from $\pi(K)$. But then $\pi^{-1}(G)$ is a hyperplane H' in $\mathbb{R}^n$, which contains o and is disjoint from $\pi^{-1}(\pi(K))$, and in particular from K. Thus, by the convexity of K, H' isolates o from K. □

The question arises whether a point can be isolated from convex sets in general. To be able to answer this question, we first prove some lemmas.

Lemma 6.7. *If $K \subseteq \mathbb{R}^n$ is convex and* $\text{int}(K) \neq \emptyset$, *then* $K \subseteq \text{cl}(\text{int}(K))$.

This statement is clearly false if $\text{int}(K) = \emptyset$.

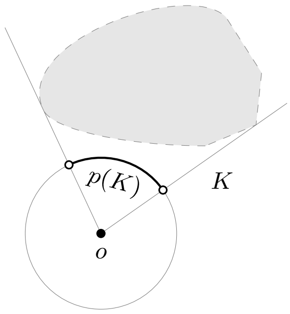


Figure 35: The image of K under the map $p(v) = \frac{v}{\|v\|}$ is an open circular arc of length at most π.

Proof. Let $p \in K$ and $q \in \operatorname{int}(K)$ be arbitrary. As $q \in \operatorname{int} K$, there is some $\varepsilon > 0$ such that the neighborhood of q of radius ε is a subset of K. But then for any point $r \in (p, q)$, the neighborhood of r of radius $\frac{\|r-p\|}{\|q-p\|}\varepsilon$ is a subset of K (Figure 37), implying that $(p, q) \subset \operatorname{int}(K)$. Thus, $p \in \operatorname{cl}(\operatorname{int}(K))$. □

Lemma 6.8. *If $K \subset \mathbb{R}^n$ is convex and $\operatorname{int}(K) = \emptyset$, then $\dim(K) < n$, or in other words, there is a hyperplane H with $K \subseteq H$.*

Proof. The proof is based on the observation that if the points $p_1, p_2, \ldots, p_{n+1}$ are affinely independent, then the interior of $\operatorname{conv}\{p_1, \ldots, p_{n+1}\}$ is not empty: indeed, if, e.g. $\frac{1}{n+1}\sum_{i=1}^{n+1} p_i$ is a boundary point of the convex hull, then by the compactness of the convex hull (Theorem 3.4) according to Corollary 2.6, there is a closed half space containing the convex hull and containing the above point in its boundary, but then by Proposition 3.5 the bounding hyperplane of this half space contains all of the p_is, which contradicts our assumption that they are affinely independent.

Now, let $p_1, \ldots, p_k$ an affinely independent point system of maximal cardinality in K. Then, by the previous observation, $k \leq n$, implying that there is a hyperplane H containing all of the points. If K has some point $p \notin H$, then it follows from Corollary 1.11 and Theorem 1.12 that $p_1, \ldots, p_k, p$ are affinely independent, which is in contradiction with the choice of the point system. Thus, $K \subseteq H$. □

Theorem 6.9 (Isolation theorem 2)**.** *Let $K \subseteq \mathbb{R}^n$ be convex with $o \notin \operatorname{int}(K)$. Then there is a hyperplane H isolating o from K.*

Proof. Assume that $\operatorname{int}(K) \neq \emptyset$. Since $\operatorname{int}(K)$ is convex (Exercise 3 from the first worksheet), by the isolation theorem there is a hyperplane H that isolates o from $\operatorname{int}(K)$. But then, since closed half spaces are closed sets, H isolates o from $\operatorname{cl}(\operatorname{int}(K))$, and thus, also from K.

Now, let $\operatorname{int}(K) = \emptyset$ and let $G = \operatorname{aff}(K)$. Then the relative interior of K is nonempty in G, and hence, there is an affine subspace G' in G for which $\dim(G') = \dim(G) - 1$, and which isolates o from K in G. But then, choosing any hyperplane H satisfying $G \cap H = G'$, H isolates o from K. □

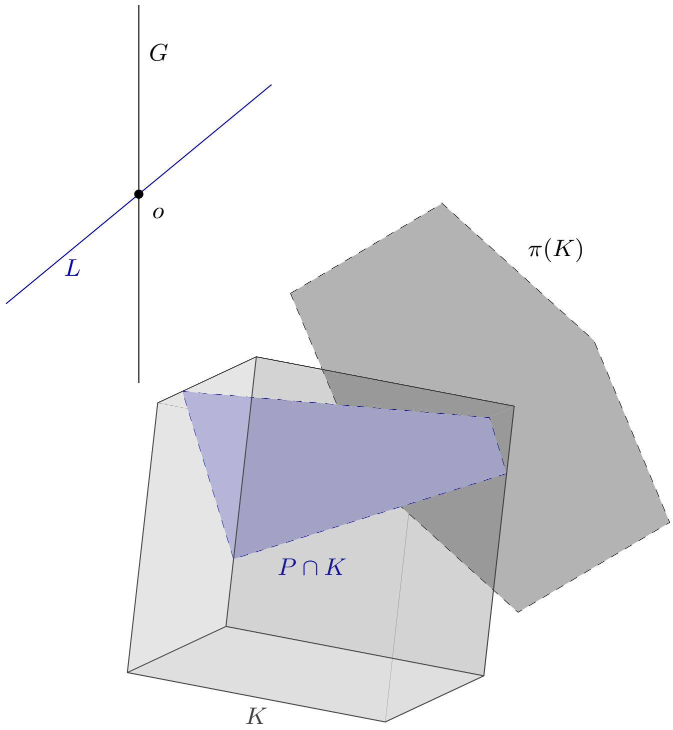


Figure 36: The open convex body $K \subseteq \mathbb{R}^n$ does not contain o. If P is an arbitrary plane containing o, then there is a line $L \subseteq P$ throught o and disjoint from $P \cap K$ ($n = 2$ case). The orthogonal projection $\pi(K)$ of K onto $H = L^\perp$ is open and disjoint from $\pi(o) = o$ therefore, by induction one can find a hyperplane G in H containing o and disjoint from $\pi(K)$. Then $\pi^{-1}(G)$ is a hyperplane in $\mathbb{R}^n$ that is disjoint from K and contains o.

Theorem 6.10. *If $K, L \subset \mathbb{R}^n$ are disjoint, convex sets, then K and L can be separated by a hyperplane.*

Proof. Let $M = K - L = K + (-1)L$. Since K and L are disjoint, $o \notin K - L$. But then, by the previous theorem, there is a hyperplane H which isolates o from M (Figure 38). In other words, there is a linear functional $f : \mathbb{R}^n \to \mathbb{R}$ satisfying $f(x) \geq 0$ for any $x \in M$. But then $M = K - L$ implies

$$\begin{aligned} 0 &\leq \inf\{f(x) | x \in M\} \\ &= \inf\{f(x) - f(y) | x \in K, y \in L\} \\ &= \inf\{f(x) | x \in K\} - \sup\{f(y) | y \in L\}. \end{aligned}$$

Let $\alpha = \inf\{f(x) | x \in K\}$. Then, according to the conditions, for any $x \in K$ we have $f(x) \geq \alpha$, and for any $x \in L$ we have $f(x) \leq \alpha$, and thus, the hyperplane $\{x | f(x) = \alpha\}$ separates K and L. $\square$

Corollary 6.11. *If $K, L \subset \mathbb{R}^n$ are disjoint, open, convex sets, then K and L can be strictly separated by a hyperplane.*

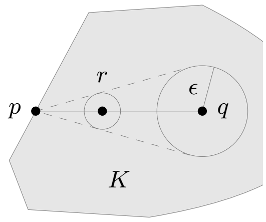


Figure 37: If $q \in \text{int}(K)$ then K contains a ball of radius $\epsilon > 0$ centered at q. For all $p \in K$ and $r \in (p,q)$ the ball of radius $\frac{\|r-p\|}{\|q-p\|}\epsilon$ centered at r is a subset of K.

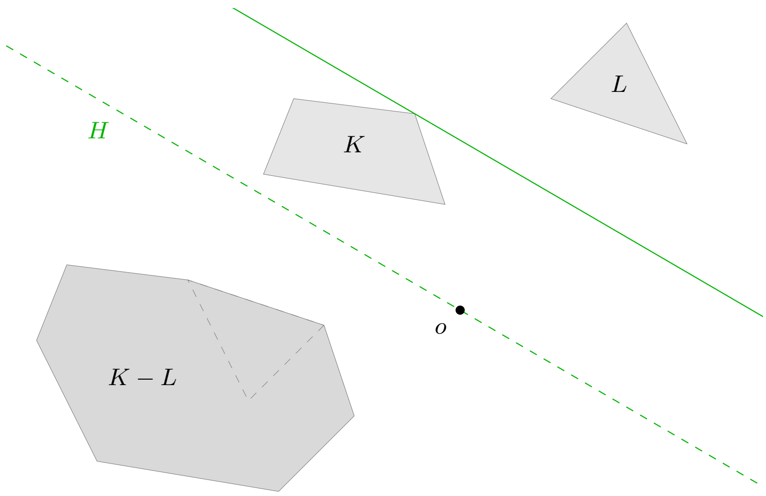


Figure 38: The hyperplane H isolates o from $K - L$, i.e., there is a linear functional f such that $f(x) \geq 0$ for all $x \in K - L$. Setting $\alpha = \inf\{f(x) | x \in K\}$, the translated hyperplane $\{x | f(x) = \alpha\}$ separates K and L.

Problem 6.12. Give an example of convex sets $K, L \subset \mathbb{R}^n$ whose interiors are disjoint, but which cannot be separated by a hyperplane.

Theorem 6.13. *Let $K, L \subset \mathbb{R}^n$ be convex sets with $\text{int}(K) \neq \emptyset$ and $\text{int}(K) \cap L = \emptyset$. Then K and L can be separated by a hyperplane.*

Proof. We have seen that if K is convex, then $\text{int}(K)$ is convex (Exercise 3 on the first worksheet). But then by Theorem 6.10, the sets $\text{int}(K)$ and L can be separated by a hyperplane. Since we learned that if $\text{int}(K) \neq \emptyset$, then $K \subset \text{cl}(\text{int}(K))$, and a hyperplane separating $\text{int}(K)$ and L separates also $\text{cl}(\text{int}(K))$ and L, the assertion follows. □

Theorem 6.14. *If $K, L \subset \mathbb{R}^n$ are disjoint, convex sets, K is compact and L is closed, then K and L can be strictly separated by a hyperplane.*

Proof. We apply the idea of Theorem 2.5. Let $x \in K$ and $y \in L$ be arbitrarily chosen points, and let $r = \|y - x\|$. Let L_0 be the set of the points of L whose distance from a point of K is at most r; in other words, let $L_0 = L \cap (K + r\mathbf{B}^n)$, where $\mathbf{B}^n$ is the closed unit ball centered at o (Figure 39). Then the distance between any points of $L \setminus L_0$ and K is greater than r, yielding that $\operatorname{dist}(K, L) = \operatorname{dist}(K, L_0)$, where $\operatorname{dist}(A, B) = \inf\{\|a - b\| | a \in A, b \in B\}$. But both K and L_0 are compact sets, and hence, there are points $x \in K$ and $y \in L$ for which $\operatorname{dist}(x, y)$ is minimal. Let H be the hyperplane bisecting the segment $[x, y]$. Then H strictly separates K and L, as otherwise there are points $x' \in K$ and $y' \in L$ for which $\|x' - y'\| < \|x - y\|$. □

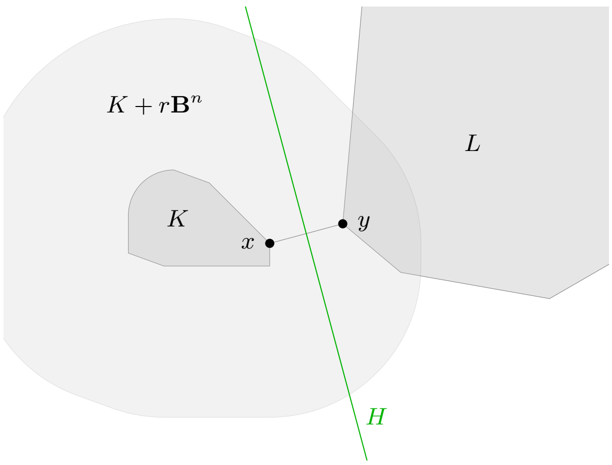


Figure 39: Choosing a sufficiently large r, the set $L_0 = L \cap (K + r\mathbf{B}^n)$ is nonempty and has the same distance to K as L. As K and L_0 are compact, there exist points $x \in K$ and $y \in L_0$ having minimal distance. The hyperplane H bisecting $[x, y]$ strictly separates K and L.

7 Faces of convex sets, extremal and exposed points, the Krein–Milman theorem

We have already seen (Corollary 2.6) that for every boundary point of a convex set there is a hyperplane through the point such that the set is contained in one of the two closed half spaces bounded by the hyperplane. This is the motivation behind the following definitions.

Definition 7.1. Let $K \subseteq \mathbb{R}^n$ be a convex set. If H is a closed half space satisfying $K \subseteq H$ and whose boundary intersects the boundary of K, we say that H is a *supporting half space* of K, and the boundary of H is a *supporting hyperplane* of K (Figure 40).

Definition 7.2. Let $K \subseteq \mathbb{R}^n$ be a closed, convex set and let H be a supporting hyperplane of K. Then the set $H \cap K$ is called a *proper face* of K (Figure 41). The empty set is called a not proper face of K. The 0-dimensional faces (consisting of only one point) are called the *exposed points* of K, and their set is denoted by $\operatorname{ex}(K)$ (Figure 42).

Our first observation implies the next remark in a natural way.

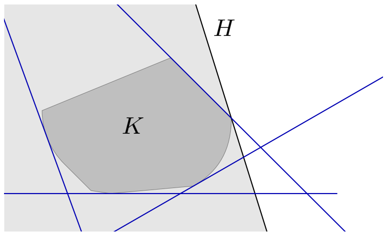


Figure 40: A supporting half space H of K and various supporting hyperplanes (blue).

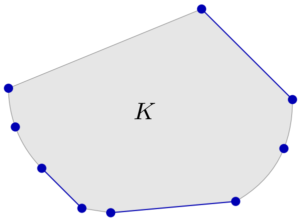


Figure 41: Proper faces of K include the marked points and closed line segments.

Remark 7.3. *If $K \subseteq \mathbb{R}^n$ is closed and convex, and $p \in \mathbb{R}^n$ is a boundary point of K, then K has a proper face F such that $p \in F$.*

Problem 7.4. Construct closed, convex sets which have no exposed points.

Proposition 7.5. *If F is a proper face of the closed, convex set $K \subseteq \mathbb{R}^n$, then F is closed and convex.*

Proof. Since every proper face F of K can be written as $F = K \cap H$, where H is a supporting hyperplane of K, and a hyperplane is closed and convex, the assertion follows from the fact that the intersection of closed, convex sets is closed and convex. □

Definition 7.6. Let $K \subseteq \mathbb{R}^n$ be closed and convex. If $p \in \operatorname{bd} K$, and for every $q, r \in K$, $p \in [q, r]$ we have $p = q$ or $p = r$, then we say that p is an *extremal point* of K. In other words, the extremal points of K are the points of K that are not relative interior points of a segment in K. The set of the extremal points of K is denoted by $\operatorname{ext}(K)$ (Figure 43).

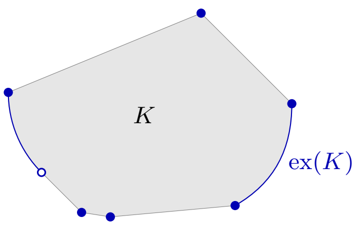


Figure 42: The set of exposed points of K.

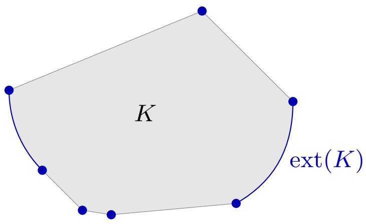
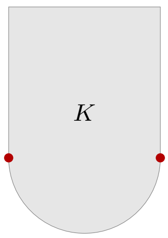


Figure 43: The set of extremal points of a closed, convex set K.

Proposition 7.7. *If $K \subseteq \mathbb{R}^n$ is closed and convex, then* $\mathrm{ex}(K) \subseteq \mathrm{ext}(K)$.

Proof. Let p be an exposed point of K. Then there is a closed half space H bounded by a hyperplane H_0 such that $K \subseteq H$ and $K \cap H_0 = \{p\}$. Assume that $q, r \in K$ and $p \in [q, r]$. Using Proposition 3.5 with $X = \{q, r\}$, we have $p \in \mathrm{conv}(X) \cap H_0 = \mathrm{conv}(X \cap H_0) \subseteq \mathrm{conv}(K \cap H_0) = \{p\}$. This implies that $X \cap H_0$ is a nonempty subset of $\{p\}$ therefore $q = p$ or $r = p$. □

Example 7.8. *Let $K \subseteq \mathbb{R}^2$ be the union of the unit square $[0, 1]^2$ and the circular region defined by the inequality $(x - 1/2)^2 + y^2 \leq 1/4$. then o and the point $(1, 0)$ are extremal points of K, but not exposed points of K. Thus, there are closed, convex sets K for which* $\mathrm{ex}(K)$ *and* $\mathrm{ext}(K)$ *do not coincide.*



Our next theorem explores the connection between extremal points and linear functionals.

Theorem 7.9. *Let $K \subseteq \mathbb{R}^n$ be a closed, convex set, and let $f : \mathbb{R}^n \to \mathbb{R}$ be a linear functional whose minimal or maximal value on K is α. Let $F = K \cap f^{-1}(\alpha)$. Then $p \in F$ is an extremal point of F if and only if it is an extremal point of K. In other words,* $\mathrm{ext}(F) = \mathrm{ext}(K) \cap f^{-1}(\alpha)$.

Before proving Theorem 7.9, we observe that if $p \in \mathrm{ex}(K)$, then there is a linear functional $f : \mathbb{R}^n \to \mathbb{R}$ which attains its minimum on K only at p. Thus, a consequence of this theorem is the containment $\mathrm{ex}(K) \subseteq \mathrm{ext}(K)$ for every closed, convex set K.

Proof. Assume that $p \in \mathrm{ext}(K)$ and $p \in F$. Then, by the definition of extremal point, for any $q, r \in K$, $p \in [q, r]$ we have $q = p$ or $r = p$. In particular, this holds also for any $q, r \in F$, implying that $p \in \mathrm{ext}(F)$.

Now, let $p \in \mathrm{ext}(F)$, and consider points $q, r \in K$ with $p \in [q, r]$. If $q \neq p$ and $r \neq p$, then for a suitable $t \in (0, 1)$, $p = tq + (1 - t)r$. But from this $\alpha = f(p) =$

$f(tq+(1-t)r) = tf(q)+(1-t)f(r)$. As $f(q), f(r) \geq \alpha$, there is equality if and only if $f(q) = f(r) = \alpha$, i.e., if $q, r \in F$. But as $p \in \text{ext}(F)$, this yields $q = p$ or $r = p$, which is a contradiction. □

Our next theorem shows an important property of extremal points.

Theorem 7.10 (Krein, Milman)**.** *Any compact, convex set $K \subset \mathbb{R}^n$ is the convex hull of its extremal points.*

Proof. We prove the statement by induction on the dimension. Assume that $K \subset \mathbb{R}$ is a compact, convex set. Then K is a closed segment, whose extremal points are its endpoints, and the segment is the convex hull of its endpoints. Thus, the assertion holds for $n = 1$.

Assume that the statement is true for any at most $(n-1)$-dimensional compact, convex set, and let K be an n-dimensional compact, convex set. Let $p \in K$ be arbitrary, and let L be an arbitrary line through p. According to our conditions, $L \cap K$ is a closed bounded segment. Let the endpoints of this segment be q and r, where these points may not be distinct from each other or p. Then, by Remark 7.3, there are faces F_q and F_r of K such that $q \in F_q$ and $r \in F_r$ (Figure 44). But as F_q and F_r are convex subsets of the boundary of K, they have no interior points, and thus, by Lemma 6.8, they are at most $(n-1)$-dimensional compact, convex sets. By the induction hypothesis, we have $q \in \operatorname{conv}\operatorname{ext}(F_q)$ and $r \in \operatorname{conv}\operatorname{ext}(F_r)$. But by the definition of face, there are linear functionals $f_q : \mathbb{R}^n \to \mathbb{R}$ and $f_r : \mathbb{R}^n \to \mathbb{R}$ attaining their minima exactly at F_q and F_r, respectively, and thus, by Theorem 7.9, the extremal points of F_q and F_r are extremal points of K. But then $p \in [q,r] \subseteq \operatorname{conv}(\operatorname{ext}(F_q) \cup \operatorname{ext}(F_r)) \subseteq \operatorname{conv}(\operatorname{ext}(K))$. □

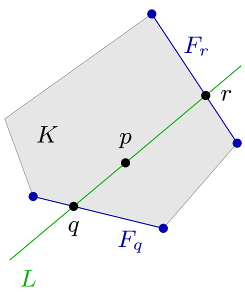


Figure 44: If p is an arbitrary point of K and L is a line through p, then $L \cap K$ is a line closed segment. Its endpoints q and r are contained in faces F_q and F_r of K, which are convex sets of lower dimension, therefore by the induction hypothesis $q \in \operatorname{conv}(\operatorname{ext}(F_q))$ and $r \in \operatorname{conv}(\operatorname{ext}(F_r))$, implying that $p \in [q,r] \subseteq \operatorname{conv}(\operatorname{ext}(K))$.

We have seen that the extremal points of a set are not necessarily exposed points. On the other hand, it is true that they are accumulation points of sequences of exposed points.

Theorem 7.11 (Straszewicz)**.** *For any compact, convex set $K \subset \mathbb{R}^n$ we have $K = \operatorname{cl}(\operatorname{conv}(\operatorname{ex}(K)))$; or in other words, K is equal to the closure of convex hull of its exposed points.*

Proof. Let $x \in \mathrm{ext}(K)$ and $\varepsilon > 0$ be arbitrary. Let us consider the compact, convex set $K_\varepsilon = \mathrm{conv}(K \setminus \mathrm{int}\, B_\varepsilon(x)) \subseteq K$, where $B_\varepsilon(x)$ denotes the closed ball of radius ε and center x. If $x \in K_\varepsilon$, then by the Carathéodory theorem it is the convex combination of at most $n+1$ points of $(K \setminus \mathrm{int}\, B_\varepsilon(x))$; that is, it is a relative interior point of a segment in K. But this contradicts the assumption that $x \in \mathrm{ext}(K)$, and thus, $x \notin K_\varepsilon$.

Note that K_ε is a compact, convex set, and thus, it can be strictly separated from x. In other words, there is a hyperplane H such that one of the closed half spaces bounded by it intersects K in a subset of $B_\varepsilon(x)$, and this half space contains x in its interior. Let L be the half line starting at x, perpendicular to H and intersecting H. For any $y \in L$ let $z(y)$ be a farthest point of K from y. Then $z(y) \in \mathrm{ex}(K)$ for any $y \in L$ (see Problem sheet 5, Exercise 4). On the other hand, if y is sufficiently far from x, then $z(y) \in B_\varepsilon(x)$. Thus $x \in \mathrm{cl}(\mathrm{ex}(K))$, from which $\mathrm{ext}(K) \subseteq \mathrm{cl}(\mathrm{ex}(K))$ (Figure 45).

By the containment relation $\mathrm{conv}(\mathrm{cl}(X)) \subseteq \mathrm{cl}(\mathrm{conv}(X))$, satisfied for any set $X \subseteq \mathbb{R}^n$, and by the Krein–Milman theorem, we have

$$K \subseteq \mathrm{conv}(\mathrm{ext}(K)) \subseteq \mathrm{conv}(\mathrm{cl}(\mathrm{ex}(K))) \subseteq \mathrm{cl}(\mathrm{conv}(\mathrm{ex}(K))) \subseteq K,$$

that is, $K = \mathrm{cl}(\mathrm{conv}(\mathrm{ex}(K)))$. □

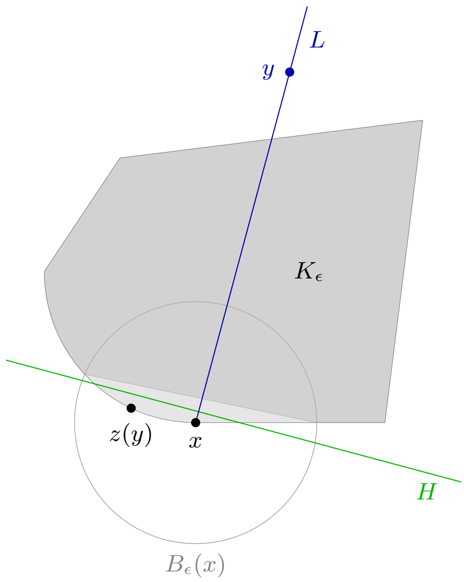


Figure 45: The set K_ϵ is the convex hull of K (light gray) with an ϵ-ball around the extremal point x removed. Since K_ϵ is compact and convex and does not contain x, it can be strictly separated from x by a hyperplane H. Let y be a point on the half line L starting at x and perpendicular to H. If y is sufficiently far from x, then the unique point $z(y) \in K$ farthest from y is in $B_\epsilon(x) \cap \mathrm{ex}(K)$.

8 Valuations and the Euler characteristic

Let us recall the following concept from our previous studies.

Definition 8.1. Let $A \subset \mathbb{R}^n$ be a set. The *indicator function* $I[A]$ of the set is the function

$$I[A](x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}.$$

We remark that for any $A, B \subset \mathbb{R}^n$, we have $I[A] \cdot I[B] = I[A \cap B]$.

Lemma 8.2 (Inclusion-exclusion formula)**.** *For any sets* $A_1, A_2, \ldots, A_k \subset \mathbb{R}^n$*,*

$$\begin{aligned} I[A_1 \cup A_2 \cup \ldots \cup A_k] &= 1 - (1 - I[A_1])(1 - I[A_2]) \ldots (1 - I[A_k]) \\ &= \sum_{j=1}^{k} (-1)^{j-1} \sum_{1 \le i_1 < i_2 < \ldots < i_j \le k} I[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_j}]. \end{aligned}$$

Proof. Let us introduce the notation $\bar{B} = \mathbb{R}^n \setminus B$ for any set $B \subseteq \mathbb{R}^n$. Observe that the first statement is equivalent to the equality

$$A_1 \cup A_2 \cup \ldots \cup A_k = \overline{\bar{A}_1 \cap \bar{A}_2 \cap \ldots \bar{A}_k},$$

which readily follows from the de Morgan identities. The second statement is a consequence of the previous remark. □

Definition 8.3. The real vector space generated by the indicator functions $I[A]$ of the compact, convex sets $A \subset \mathbb{R}^n$ is called the *algebra of compact, convex sets*, and is denoted by $\mathcal{K}(\mathbb{R}^n)$. The real vector space generated by the indicator functions $I[A]$ of the closed, convex sets $A \subset \mathbb{R}^n$ is called the *algebra of closed, convex sets*, and is denoted by $\mathcal{C}(\mathbb{R}^n)$.

Remark 8.4. *An arbitrary element of* $\mathcal{K}(\mathbb{R}^n)$ *can be written as* $\sum_{i=1}^{k} \alpha_i I[A_i]$*, where* $\alpha_i \in \mathbb{R}$*, and the sets* $A_i \subset \mathbb{R}^n$ *are compact and convex. Observe that if* $A, B \subset \mathbb{R}^n$ *are compact, convex sets, then* $A \cap B$ *is also compact and convex, implying that the product of two elements of* $\mathcal{K}(\mathbb{R}^n)$ *is also an element of* $\mathcal{K}(\mathbb{R}^n)$*. Thus, the set* $\mathcal{K}(\mathbb{R}^n)$ *is indeed an algebra over* $\mathbb{R}$*. A similar observation can be made about the algebra* $\mathcal{C}(\mathbb{R}^n)$*.*

Definition 8.5. A linear map $\mathcal{K}(\mathbb{R}^n) \to \mathbb{R}$ or $\mathcal{C}(\mathbb{R}^n) \to \mathbb{R}$ is called a *valuation*.

The main goal of this lecture is the proof of the next theorem.

Theorem 8.6. *There is a unique valuation* $\chi : \mathcal{C}(\mathbb{R}^n) \to \mathbb{R}$ *satisfying* $\chi(I[A]) = 1$ *for all nonempty, closed, convex sets* $A \subset \mathbb{R}^n$*.*

This valuation is called the *Euler characteristic* induced by the algebra of closed, convex sets. Theorem 8.6 was first proved by H. Hadwiger.

Proof. Note that by the linearity of χ, it can be uniquely extended to every element of $\mathcal{C}(\mathbb{R}^n)$, implying that χ is unique. We need to show that χ exists. We first define this valuation on the elements of $\mathcal{K}(\mathbb{R}^n)$ by induction on the dimension.

Assume that $n = 0$. Then any function $f \in \mathcal{K}(\mathbb{R}^0)$ can be written as $f = \alpha I[o]$ for some $\alpha \in \mathbb{R}$. Thus, $\chi(f) = \alpha$ satisfies the conditions of the theorem.

Let $n > 0$. For any $x \in \mathbb{R}^n$, let $p(x)$ denote the last coordinate of x, and for any $t \in \mathbb{R}$, define the hyperplane

$$H_t = \{x \in \mathbb{R}^n | p(x) = t\} .$$

This hyperplane can be identified with $\mathbb{R}^{n-1}$, and thus, there is a (unique) valuation χ_t on it satisfying the conditions of the theorem. For any $f \in \mathcal{K}(\mathbb{R}^n)$, let f_t denote the restriction of f onto H_t. Then, if $f = \sum_{i=1}^{k} \alpha_i I[A_i]$, where $\alpha_i \in \mathbb{R}$ and the A_is are compact, convex sets, then

$$f_t = \sum_{i=1}^{k} \alpha_i I[A_i \cap H_t],$$

and hence, by $f_t \in \mathcal{K}(H_t)$, we have

$$\chi_t(f_t) = \sum_{\substack{i \\ A_i \cap H_t \neq \emptyset}} \alpha_i.$$

Consider the limit

$$\lim_{\varepsilon \to 0^+} \chi_{t-\varepsilon}(f_{t-\varepsilon}).$$

Note that this limit is equal to $\chi_t(f_t)$ if and only if for any sufficiently small $\varepsilon > 0$ and for every value of i, $A_i \cap H_t \neq \emptyset$ implies $A_i \cap H_{t-\varepsilon} \neq \emptyset$ (Figure 46).

In general, we have that $\lim\limits_{\varepsilon \to 0^+} \chi_{t-\varepsilon}(f_{t-\varepsilon})$ is equal to the sum of the α_is for which, for any small $\varepsilon > 0$, we have $A_i \cap H_{t-\varepsilon} \neq \emptyset$. That is, the limit is $\chi_t(f_t)$ unless t is the minimum of the orthogonal projection p on a set A_i. Thus, for any function f, the limit differs from $\chi_t(f_t)$ only for finitely many values of t. Based on this, we define the function χ as

$$\chi(f) = \sum_{t \in \mathbb{R}} \left(\chi_t(f_t) - \lim_{\varepsilon \to 0^+} \chi_{t-\varepsilon}(f_{t-\varepsilon}) \right) .$$

Consider the functions $f, g \in \mathcal{K}(\mathbb{R}^n)$ and numbers $\alpha, \beta \in \mathbb{R}$. Since the valuation χ_t, and the operation of taking limit, are linear, it follows that $\chi(\alpha f + \beta g) = \alpha\chi(f) + \beta\chi(g)$. Furthermore, if $A \subset \mathbb{R}^n$ is a nonempty, compact, convex set, then

$$\chi_t(I[A \cap H_t]) - \lim_{\varepsilon \to 0^+} \chi_{t-\varepsilon}(I[A \cap H_{t-\varepsilon}]) = \begin{cases} 1, & \text{if } \min\limits_{x \in A} p(x) = t, \\ 0, & \text{otherwise.} \end{cases}$$

As the minimum is uniquely defined on A, we have $\chi(I[A]) = 1$.

Now we extend χ to $\mathcal{C}(\mathbb{R}^n)$. Using the standard notation $B_\rho(o) = \{x \in \mathbb{R}^n | \|x\| \leq \rho\}$, if $f \in \mathcal{C}(\mathbb{R}^n)$, let

$$\chi(f) = \lim_{\rho \to \infty} f \cdot I[B_\rho(o)].$$

Then χ clearly satisfies the requirements. □

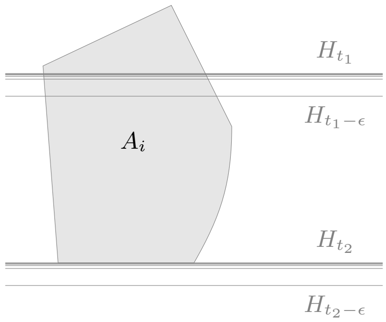
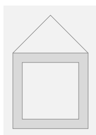


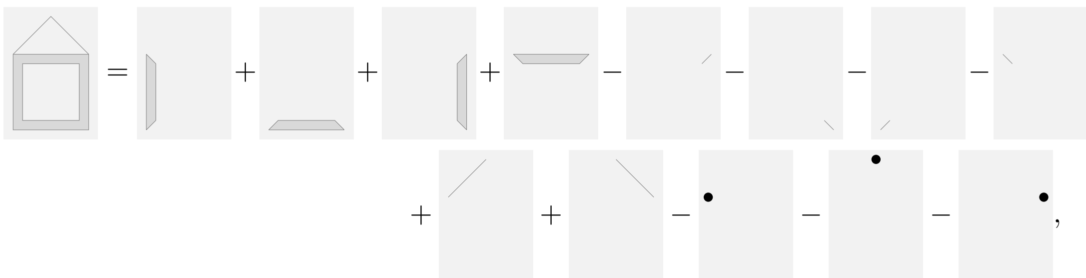
Figure 46: The sets $A_i \cap H_{t_1}$ and $A_i \cap H_{t_2}$ are nonempty, compact and convex, therefore $\chi_{t_1}(I(A_i \cap H_{t_1})) = \chi_{t_2}(I(A_i \cap H_{t_2})) = 1$. For small $\epsilon > 0$ the set $A_i \cap H_{t_1-\epsilon}$ is also nonempty, compact and convex, therefore $\chi_{t_1-\epsilon}(I(A_i \cap H_{t_1-\epsilon})) = 1$. On the other hand, $A_i \cap H_{t_2-\epsilon} = \emptyset$ for $\epsilon > 0$, therefore $\chi_{t_1-\epsilon}(I(A_i \cap H_{t_1-\epsilon})) = 0$, implying that $\chi_t(I(A_i \cap H_t))$ jumps by one at t_2, the minimum of the last coordinate over A_i.

If $A \subset \mathbb{R}^n$ is a set such that $I[A] \in \mathcal{C}(\mathbb{R}^n)$, then, instead of $\chi(I[A])$, we use the notation $\chi(A)$. We call this quantity the *Euler characteristic of A*. We remark that Euler characteristic can be also defined in a more general setting, for the so-called *CW complexes*. Nevertheless, the discussion of these complexes is outside the scope of this course.

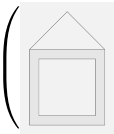
Example 8.7. *Consider the following shape in the plane (with a light background added for reference):*



A possible way to express its indicator function as a linear combination of indicator functions of compact convex sets is



and from this we infer

$\chi\Bigg($  $\Bigg)$

$$= 1+1+1+1-1-1-1-1 \\ +1+1-1-1-1 \\ = -1.$$

In the proof of the previous theorem, we proved also the following lemma.

Lemma 8.8. *Let $A \subset \mathbb{R}^n$ be a set such that $I[A] \in \mathcal{K}(\mathbb{R}^n)$. Let $t \in \mathbb{R}$, and let H_t be the set of the points $x = (x_1, \ldots, x_n)$ with $x_n = t$. Then $I[A \cap H_t] \in \mathcal{K}(\mathbb{R}^n)$, and*

$$\chi(A) = \sum_{t \in \mathbb{R}} \left(\chi(A \cap H_t) - \lim_{\varepsilon \to 0^+} \chi(A \cap H_{t-\varepsilon}) \right).$$

The last lemma is the consequence of Lemma 8.2 and Theorem 8.6.

Lemma 8.9. *Let $A_1, A_2, \ldots, A_k \subset \mathbb{R}^n$ be sets such that $I[A_i] \in \mathcal{K}(\mathbb{R}^n)$ for any $i = 1, 2, \ldots, k$. Then*

$$\chi(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{j=1}^{k} (-1)^{j-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq k} \chi(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_j}).$$

9 Convex polytopes and polyhedral sets

Our next topic is the theory of convex polytopes. Our main concept is as follows.

Definition 9.1. The convex hull of finitely many points in $\mathbb{R}^n$ is called a *convex polytope*, or shortly, *polytope* (Figure 47). If $P \subset \mathbb{R}^n$ is a convex polytope, then the set $\{x_1, x_2, \ldots, x_k\} \subset \mathbb{R}^n$ is a *minimal representation* of P, if

(i) $P = \operatorname{conv}\{x_1, x_2, \ldots, x_k\}$, and

(ii) for any index i, we have $x_i \notin \operatorname{conv}\{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k\}$.

Let us observe that every convex polytope has a minimal representation, which can be obtained by removing redundant points one by one from any represention. It is worth noting that the exposed points (that is, 0-dimensional faces) of a convex polytope are usually called *vertices*, and the $(n-1)$-dimensional faces of a convex polytope are called *facets*.

Theorem 9.2. *Let $M = \{x_1, \ldots, x_k\} \subset \mathbb{R}^n$ be a minimal representation of the convex polytope P. Then the following are equivalent:*

(i) $x \in M$,

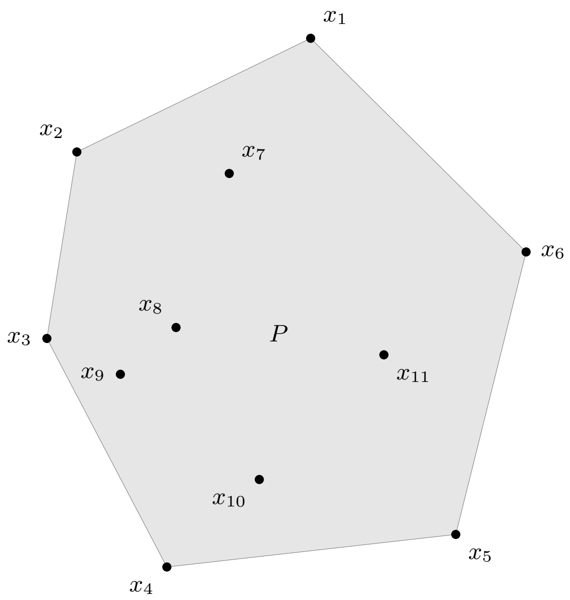


Figure 47: The set $P = \operatorname{conv}\{x_1, \ldots, x_{11}\}$ is a convex polytope in $\mathbb{R}^2$. Its minimal representation is the set $\{x_1, \ldots, x_6\}$. The exposed points (vertices) are $x_1, \ldots, x_6$, and its facets are $[x_1, x_2], \ldots, [x_5, x_6], [x_6, x_1]$.

(ii) $x \in \operatorname{ex}(P)$,

(iii) $x \in \operatorname{ext}(P)$.

Proof. (i) $\Longrightarrow$ (ii): Assume that $x \in M$. Then $x \notin \operatorname{conv}(M \setminus \{x\})$. Since $\operatorname{conv}(M \setminus \{x\})$ is compact and convex, there is a hyperplane H that strictly separates it from x. Let H_0 be the hyperplane parallel to H and containing x (see Figure 48). Then $H_0 \cap M = \{x\}$ and H_0 is a supporting hyperplane of $P = \operatorname{conv}(M)$. By Proposition 3.5, then $H_0 \cap P = H_0 \cap \operatorname{conv}(M) = \operatorname{conv}(H_0 \cap M) = \{x\}$, and hence, x is a vertex of P.

(ii) $\Longrightarrow$ (iii): By Proposition 7.7, for any closed, convex set K we have $\operatorname{ex}(K) \subseteq \operatorname{ext}(K)$. As M is compact, so is its convex hull by Theorem 3.4.

(iii) $\Longrightarrow$ (i): Let $x \in \operatorname{ext}(P)$. Now, if $x \in \operatorname{conv}(M \setminus \{x\})$ was true, then x could be written as a convex combination of points from $M \setminus \{x\}$. Choosing a minimal number of such points one can show that then x could be written as a relative interior point of a segment in P, which would contradict the condition that $x \in \operatorname{ext}(P)$. □

Corollary 9.3. *Every convex polytope has a unique minimal representation.*

Remark 9.4. *By Proposition 3.5, if H is a supporting hyperplane of the convex set $\operatorname{conv}(X)$, then $H \cap \operatorname{conv}(X) = \operatorname{conv}(H \cap X)$. From this it follows that every face of a convex polytope is a convex polytope, and also that every convex polytope has only finitely many faces.*

The next two statements hold for the faces of every compact, convex sets.

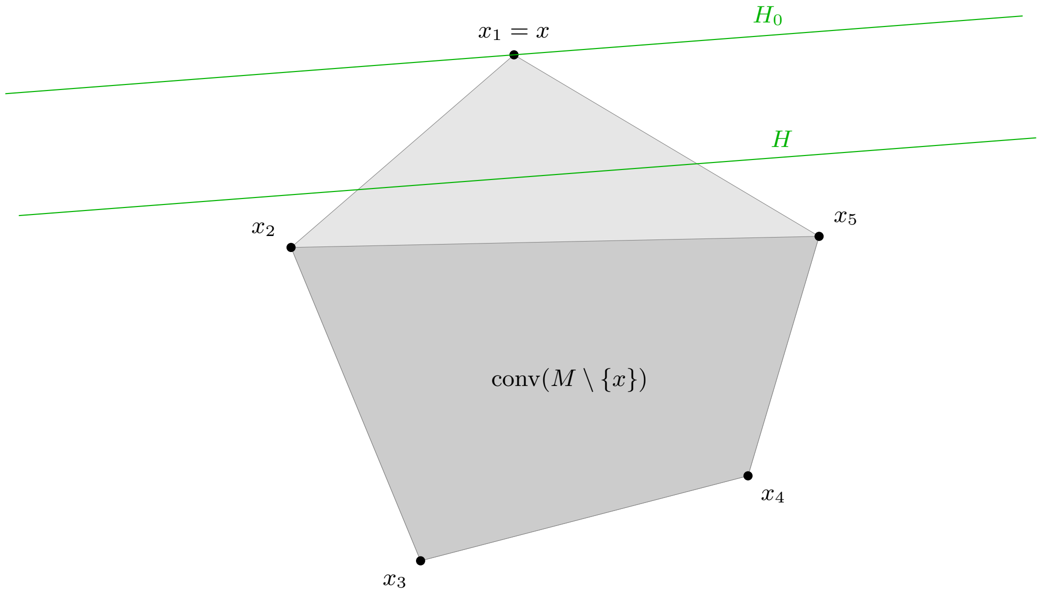


Figure 48: Since $M = \{x_1, \dots, x_k\}$ is a minimal representation of the polytope P, for $x \in M$ we have $x \notin \operatorname{conv}(M \setminus \{x\})$, therefore a hyperplane H strictly separates $\{x\}$ and $\operatorname{conv}(M \setminus \{x\})$. If H_0 is the translate of H that contains x, then $H_0 \cap P = \{x\}$.

Proposition 9.5. *If $K \subset \mathbb{R}^n$ is a nonempty, compact, convex set, and $F_1, \dots, F_m$ are faces of K, then $F = \bigcap_{i=1}^m F_i$ is a face of K.*

Proof. If $F = \emptyset$, then F is a face of K, and thus, we may assume that $F \neq \emptyset$, which implies that for every i, F_i is a proper face of K. Without loss of generality, we may assume that $o \in F$. Since F_i is a proper face of K, there is a linear functional $f_i : \mathbb{R}^n \to \mathbb{R}$ satisfying $f_i(x) \geq 0$ for all $x \in K$, and for which $f(x) = 0$ for some $x \in K$ if and only if $x \in F_i$. Now, let $f = \sum_{i=1}^m f_i$. This function f is a linear functional, and if $x \in K$, then $f(x) \geq 0$. Assume that $x \in K$ and $f(x) = 0$. From this, $\sum_{i=1}^m f_i(x) = 0$, but since $f_i(x) \geq 0$ for any value of i, this is satisfied if and only if $f_i(x) = 0$ for all values of i, or in other words, if $x \in F$. Thus, F is a face of K. □

Proposition 9.6. *Let $S_2 \subseteq S_1 \subset \mathbb{R}^n$ be compact, convex sets. If F is a face of S_1, then $F \cap S_2$ is a face of S_2.*

Proof. If $F \cap S_2 = \emptyset$, then it is clearly a face of S_2. Assume that $F \cap S_2 \neq \emptyset$, which implies that F is a proper face of S_1. Let H be a supporting hyperplane of S_1 satisfying $H \cap S_1 = F$. Then H also supports S_2, and $H \cap S_2 = (H \cap S_1) \cap S_2 = F \cap S_2$, implying that $F \cap S_2$ is a face of S_2. □

Our next proposition, which, in some sense, is the converse of the previous one, holds only for convex polytopes.

Proposition 9.7. *Let F_1 be a proper face of a convex polytope P, and let F_2 be a face of F_1. Then F_2 is a face of P.*

Proof. If $F_2 = \emptyset$, then the statement holds, and hence, we may assume that F_2 is a proper face of F_1. According to our conditions, P has a supporting hyperplane H in $\mathbb{R}^n$ satisfying $P \cap H = F_1$, and if F_2 is a proper face of F_1, then there is a 'supporting hyperplane' G of F_2 in H satisfying $G \cap F_1 = F_2$. Observe that $\dim G = n - 2$. As P is a convex polytope, only finitely many vertices of P are not elements of H, and thus, H can be rotated around G with a sufficiently small angle in a suitable direction such that the hyperplane H' obtained by this rotation is a supporting hyperplane of P, and, from amongst the vertices of P, it contains only those in F_2 (Figure 49). But from this, it follows that $H' \cap P = F_2$, yielding that F_2 is a face of P. $\square$

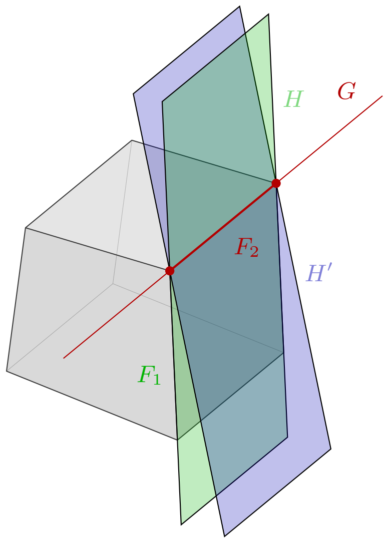


Figure 49: Since F_1 is a face of the convex polytope P, there is a supporting hyperplane $H \subseteq \mathbb{R}^n$ such that $H \cap P = F_1$. Similarly, there is an affine subspace $G \subseteq H$ of dimension $n-2$ such that $G \cap F_1 = F_2$. Rotating H by small angle results in a hyperplane H' satisfying $H \cap P = F_2$.

Problem 9.8. Construct a compact, convex set $K \subseteq \mathbb{R}^n$ with the property that it has a proper face F_1, and F_1 has a proper face F_2 such that F_2 is not a face of K.

We have seen that every compact, convex set can be obtained as the intersection of closed half spaces. Now we show that a convex polytope is the intersection of finitely many closed half spaces (namely those defined by its facets).

Definition 9.9. The intersection of finitely many closed half spaces is called a *polyhedral set* (Figure 50).

Theorem 9.10. *Every convex polytope is a bounded polyhedral set.*

Proof. Let $P \subset \mathbb{R}^n$ be a convex polytope. As P is compact, it is sufficient to prove that it is a polyhedral set. Without loss of generality, assume that $\dim P = n$, as every

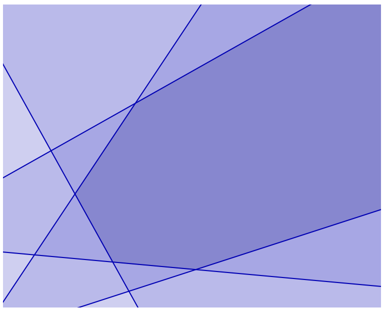


Figure 50: A polyhedral set is the intersection of finitely many closed half spaces.

hyperplane is obtained as the intersection of the two closed half spaces it generates, and every affine subspace is obtained as the intersection of finitely many hyperplanes.

Let $M = \{x_1, \ldots, x_k\}$ be a minimal representation of P. Let the facets of P be $F_1, \ldots, F_m$, and denote by H_i and H_i^+ the supporting hyperplane and the closed supporting half space defined by F_i, respectively. Then for any index i, we have $P \cap H_i = F_i$ and $P \subset H_i^+$. We show that $P = \bigcap_{i=1}^m H_i^+$.

Cearly, $P \subseteq \bigcap_{i=1}^m H_i^+$, and thus, by contradiction, we suppose that there is a point $x \in \left(\bigcap_{i=1}^m H_i^+\right) \setminus P$. Now, let $D = \bigcup \operatorname{aff}(\{x\} \cup C)$, where C runs over the family of the subsets of M of cardinality at most $(n-1)$. Then D is the union of finitely many affine subspaces of dimension at most $(n-1)$, and thus, we can choose a point $y \in \operatorname{int}(P)$ with $y \notin D$. But then, by $x \notin P$, the segment $[x, y]$ intersects the boundary of P, that is, there is a point $z \in (x, y)$ with $z \in \operatorname{bd}(P)$. We will show that z lies on a facet of P, but it does not lie on any lower dimensional face of P.

Assume that $z \in F$ for some j-dimensional face of P, where $0 \leq j \leq n-2$. Then, by Carathéodory's theorem, z is contained in the convex hull of at most $(n-1)$ points of M, implying $\operatorname{aff}\{x, z\} \in D$, which contradicts the assumption that $y \notin D$. By Corollary 2.6, any boundary point of a compact, convex set is a point of a supporting hyperplane of the set, and thus, a point of a proper face of the set. Thus, by exclusion, z is a point of a facet F_i of P. But from this, by $y \in \operatorname{int} P \subset \operatorname{int} H_i^+$, we obtain $x \notin H_i$, which contradicts our choice of y. This yields that $P = \bigcap_{i=1}^m H_i^+$. □

Corollary 9.11. *The boundary of every n-dimensional convex polytope $P \subset \mathbb{R}^n$ is the union of the facets of P.*

Theorem 9.12. *Every bounded polyhedral set is a convex polytope.*

Proof. Every bounded polyhedral set $P \subset \mathbb{R}^n$ is a compact, convex set. Thus, by the Krein–Milman theorem, it is sufficient to show that P has finitely many extremal points. We prove this by induction on the dimension n. If $n = 1$, then every compact, convex set (in particular, P) is a closed segment with two extremal points, the endpoints of the segment. Thus, for $n = 1$ the statement holds. Now, let P be an n-dimensional polyhedral set, and let $H_1, \ldots, H_k$ be the hyperplanes bounding the closed half spaces defining P.

Let $x \in \text{ext}(P)$. If $x \in P$ and $x \notin H_i$ for any i, then, by the continuity of linear functionals, $x \in \text{int}(P)$, implying $x \notin \text{ext}(P)$. Thus, we can assume that $x \in H_i$ for some value of i. By Theorem 7.9, for any closed, convex set K and any supporting hyperplane H of K, we have $\text{ext}(K) \cap H = \text{ext}(K \cap H)$. This yields that $\text{ext}(H_i \cap P) = \text{ext}(P) \cap H_i$. But, by the induction hypothesis, $|\text{ext}(H_i \cap P)| < \infty$, implying $|\text{ext}(P)| \leq \sum_{i=1}^{m} |\text{ext}(H_i \cap P)| < \infty$. □

10 Face structures of polytopes and Euler characteristic

Let us recall the definition of algebraic lattice.

Definition 10.1. Let $(A, \leq)$ be a partially ordered set. If, for any $a_1, a_2, \ldots, a_k \in A$ there is a $c \in A$ such that $c \leq a_i$ for every value of i, and if $d \in A$, $d \leq a_i$ for every i implies that $d \leq c$, then we say that c is the *infimum* of $a_1, \ldots, a_k$. One can define the *supremum* of $a_1, \ldots, a_k$ similarly. If for any $a, b \in A$, a and b has an infimum and a supremum, we say that $(A, \leq)$ is an (algebraic) *lattice*.

Definition 10.2. Assume that $(A \leq)$ is a lattice with a minimal element, denoted by 0, that is, assume that there is an element $0 \in A$ such that $0 \leq a$ for all $a \in A$. We say that $a \in A$, $a \neq 0$ is an *atom*, if $b \in A$, and $b \leq a$ implies $b = a$ or $b = 0$. We say that $(A, \leq)$ is *atomic*, if for every $b \in A$, $b \neq 0$ there is an atom $a \in A$ satisfying $a \leq b$. We say that $(A, \leq)$ is *atomistic*, if every element of A is the supremum of some atoms in A.

Example 10.3. *Let n be a positive integer. The set of divisors of n, partially ordered by divisibility, is a lattice. Its minimal element is 1, and its atoms are the prime divisors of n. It is atomic since every number other than 1 has a prime divisor. The lattice is atomistic iff n is square-free (see Figures 51 and 52).*

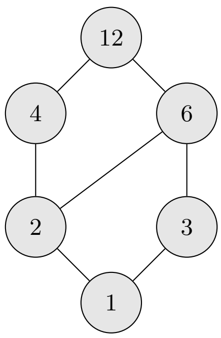


Figure 51: The Hasse diagram of the lattice of divisors of $12 = 2^2 \cdot 3$. The infimum of 4 and 6 is 2 (greatest common divisor), while the supremum of 2, 3 and 4 is 12 (least common multiple). This lattice is atomic but not atomistic, since 4 is not the supremum of any set of atoms.

Theorem 10.4. *Let $P \subset \mathbb{R}^n$ be an n-dimensional convex polytope, and let $\mathcal{F}$ the family consisting of the faces of P (including the empty set), and also P. Then $\mathcal{F}$ is a lattice*

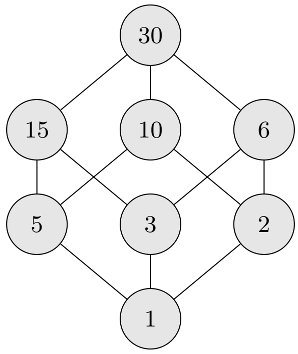


Figure 52: The Hasse diagram of the lattice of divisors of $30 = 2 \cdot 3 \cdot 5$. The minimal element is 1, the atoms are 2, 3 and 5 (prime factors), and every divisor is the least common multiple of some of the atoms, i.e., this lattice is atomistic.

with respect to the partial order defined by the containment relation. This lattice is atomic and atomistic, and its atoms are the vertices of P.

Proof. Let $F \in \mathcal{F}$. Then the infimum and the supremum of $\emptyset$ and F is $\emptyset$ and F, respectively, and the infimum and the supremum of P and F are F and P, respectively. Now, let F_1 and F_2 be proper faces of P. We have seen that $F = F_1 \cap F_2$ is a face of P. Clearly, for any $F' \in \mathcal{F}$ with $F' \subseteq F_1$ and $F' \subseteq F_2$, we have $F' \subseteq F$, and thus, F is the infimum of F_1 and F_2.

We show that F_1 and F_2 has a supremum. Indeed, if there is no proper face of P that contains both F_1 and F_2, then, clearly, P is the supremum of F_1 and F_2. If there is a proper face containing $F_1 \cup F_2$, then let F denote the intersection of all the faces satisfying this property. As F is a face of P, we have that F is the supremum of F_1 and F_2.

We have shown that $\mathcal{F}$ is a lattice. The minimal element of this lattice is $\emptyset$, and the singleton faces, i.e. the vertices, are its atoms. By the theorem of Straszewicz, every convex polytope has vertices. Furthermore, as the proper faces of P are convex polytopes, every face has vertices, yielding that the atoms are exactly the vertices of P, and $\mathcal{F}$ is atomic. On the other hand, every face is the supremum of the verties contained in the face, and thus, $\mathcal{F}$ is atomistic. □

Definition 10.5. The lattice assigned to the n-dimensional convex polytope P in Theorem 10.4 is called the *face lattice* of P.

We continue with the properties of the Euler characteristics of convex polytopes.

Lemma 10.6. *Let $P \subset \mathbb{R}^n$ be an n-dimensional (convex) polytope. Then*

$$\chi(\operatorname{bd} P) = 1 + (-1)^{n-1}, \quad \textit{and} \quad \chi(\operatorname{int} P) = (-1)^n.$$

Proof. By Corollary 9.11 $\operatorname{bd} P$ is the union of the facets of P, and thus, by Lemma 8.2 $I[\operatorname{bd} P] \in \mathcal{K}(\mathbb{R}^n)$ and thus, $\chi(\operatorname{bd} P)$ exists. We prove the first equality by induction.

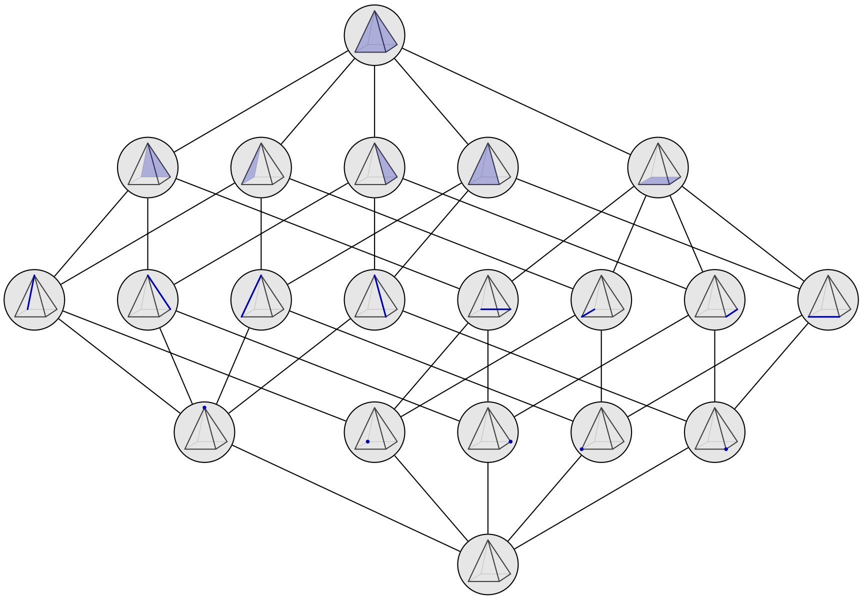


Figure 53: The Hasse diagram of the face lattice of a square pyramid.

If $n = 1$, then P is a closed segment, for which the assertion readily follows. Assume that P is an n-dimensional polytope, and also that the statement holds for $(n-1)$-dimensional polytopes. We use the notation of Lemma 8.8. By the lemma,

$$\chi(\operatorname{bd} P) = \sum_{t\in\mathbb{R}} \left(\chi(H_t \cap \operatorname{bd} P) - \lim_{\varepsilon\to 0^+} \chi(H_{t-\varepsilon} \cap \operatorname{bd} P) \right).$$

Let $t_{\min} = \min_{x\in P} x_n$ and $t_{\max} = \max_{x\in P} x_n$, where $x = (x_1, \ldots, x_n)$. Then, for every $t_{\min} < t < t_{\max}$, the set $P \cap H_t$ is an $(n-1)$-dimensional polytope, and thus, by the induction hypothesis, $\chi(H_t \cap \operatorname{bd} P) = \chi(\operatorname{bd}(H_t \cap P)) = 1 + (-1)^{n-2}$ (Figure 54). If $t = t_{\min}$ or $t = t_{\max}$, then $H_t \cap \operatorname{bd} P$ is a face of the polytope, and thus, $\chi(H_t \cap \operatorname{bd} P) = 1$. Furthermore, if $t > t_{\max}$ or $t < t_{\min}$, then $\chi(H_t \cap \operatorname{bd} P) = 0$. Summing up:

$$\chi(\operatorname{bd} P) = 1 - (1 + (-1)^{n-2}) + 1 = 1 + (-1)^{n-1}.$$

Finally, by $I[\operatorname{int} P] = I[P] - I[\operatorname{bd} P]$, we have

$$\chi(\operatorname{int} P) = 1 - (1 + (-1)^{n-1}) = (-1)^n.$$

□

Definition 10.7. Let $P \subset \mathbb{R}^n$ be an n-dimensional convex polytope. If $i = 0, 1, \ldots, n-1$, let $f_i(P)$ denote the number of the i-dimensional faces of P. Then the vector $f(P) = (f_0(P), f_1(P), \ldots, f_{n-1}(P), 1) \in \mathbb{R}^{n+1}$ is called the *f-vector* of P.

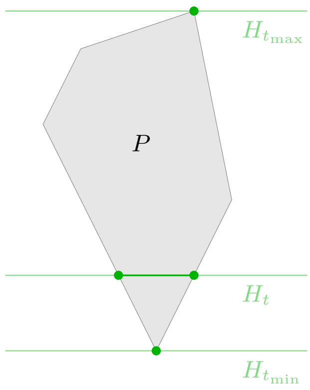


Figure 54: $\chi(H_{t_{\min}} \cap \operatorname{bd} P) = \chi(H_{t_{\max}} \cap \operatorname{bd} P) = 1$, since both intersections are convex polytopes. For $t_{\min} < t < t_{\max}$, the intersection $H_t \cap \operatorname{bd} P$ is the relative boundary of $H_t \cap P$, and $H_t \cap P$ is a convex polytope of dimension $n-1$. By the induction hypothesis, $\chi(H_t \cap \operatorname{bd} P) = 1 + (-1)^{n-2}$.

We remark that the last coordinate is the consequence of the convention, often appearing in the literature, which regards P as an n-dimensional face of itself.

Example 10.8.

(i) The f-vector of a convex n-gon is $(n, n, 1)$.

(ii) The f-vector of the cube is $(8, 12, 6, 1)$.

(iii) The f-vector of an n-dimensional simplex is $\left(n+1, \binom{n+1}{2}, \binom{n+1}{3}, \ldots, \binom{n+1}{n+1}\right)$.

To prove our next theorem we need a lemma, with respect to which we should clarify that the relative interiors of singletons (i.e. 0-dimensional affine subspaces) are themselves.

Lemma 10.9. *Let $P \subset \mathbb{R}^n$ be an n-dimensional polytope and let $x \in \operatorname{bd}(P)$ be arbitrary. Then there is a unique face of P containing x in its relative interior.*

Proof. Let F be the intersection of the faces containing x. Since P has only finitely many faces, and the intersection of finitely many faces is a face, it follows that F is a face of P. As $x \in F$, therefore F is a proper face. We show that $x \in \operatorname{relint}(F)$, and that F is the only face of P with this property.

Assume that $x \in \operatorname{relbd}(F)$. Since F is a convex polytope, F has a face F' containing x. But then Proposition 9.7 implies that F' is a proper face of P, and thus we have found a face F' containing x with $F \not\subseteq F'$, which contradicts the definition of F. Thus, $x \in \operatorname{relint}(F)$.

For contradiction, let $F' \neq F$ be a proper face of P satisfying $x \in \operatorname{relint}(F')$. Then, by the definition of F, we have $F \subset F'$. On the other hand, since F is a face of P, there is a hyperplane H supporting P with $H \cap P = F$. This hyperplane supports also the convex polytope F' in F, implying that F is a proper face of F'. Thus, $x \in F \subset \operatorname{relbd}(F')$; a contradiction. $\square$

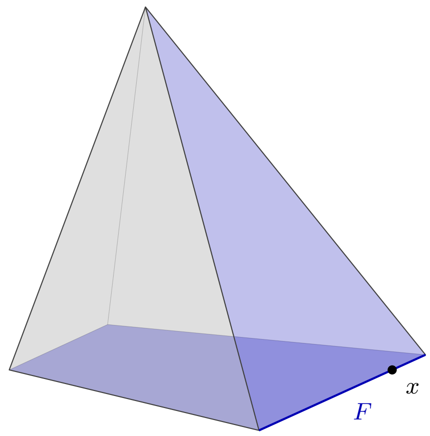


Figure 55: The point x lies in the boundary of the convex polytope P, and F is the intersection of the faces containing x. It follows that x is in the relative interior of F and no other face contains x in its relative interior.

Theorem 10.10 (Euler)**.** *Let $P \subset \mathbb{R}^n$ be an n-dimensional convex polytope. Then*

$$\sum_{i=0}^{n-1}(-1)^i f_i(P) = 1 + (-1)^{n-1}.$$

Proof. Lemma 10.9 implies that $I[P] = \sum_F I[\text{relint} F]$, where the summation is taken over all nonempty faces of P, and P itself. Applying the valuation χ to both sides of this equation, the statement follows from Lemma 10.6. □

11 Polarity

From now on, we denote by $B_r(x)$ the closed ball of radius r and center x. The main concept of this lecture is the following.

Definition 11.1. Let $A \subseteq \mathbb{R}^n$ be a nonempty set. Then the *polar of* A is the set

$$A^* = \{y \in \mathbb{R}^n | \langle x, y\rangle \leq 1 \text{ for every } x \in A\}.$$

Example 11.2.

(i) $\{o\}^* = \mathbb{R}^n$,

(ii) *If $x \neq o$, then $\{x\}^*$ is the closed half space, containing o, whose boundary is perpendicular to x and its distance from o is $\frac{1}{\|x\|}$.*

(iii) *For any $r > 0$, $B_r(o)^* = B_{\frac{1}{r}}(o)$. This statement readily follows from the previous example, since the intersection of the closed half spaces, containing o, whose distance from o is $\frac{1}{r}$ coincides with $B_{\frac{1}{r}}(o)$.*

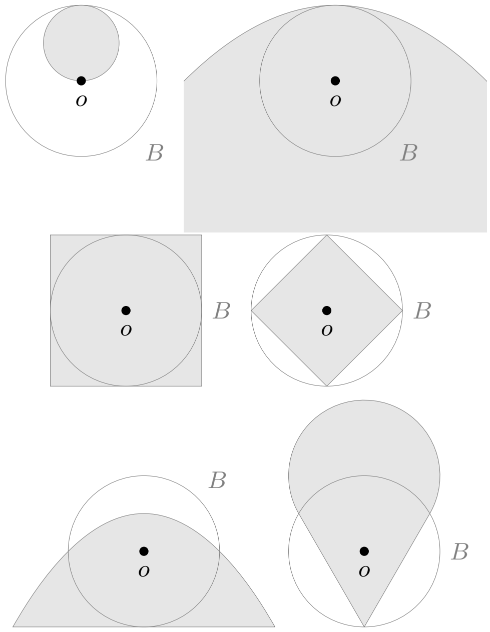


Figure 56: Sets and their polars.

The next theorem summarizes some simple properties of polarity.

Theorem 11.3.

(i) *For any set* $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$, *we have* $A^* = \bigcap_{a \in A} \{a\}^*$.

(ii) *For any nonempty sets* $A_i \subseteq \mathbb{R}^n$, $i \in I$, *we have* $\left(\bigcup_{i \in I} A_i\right)^* = \bigcap_{i \in I} A_i^*$.

(iii) *For any* $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$, *the set* A^* *is a closed, convex set containing* o.

(iv) *If* $A_1 \subseteq A_2 \subseteq \mathbb{R}^n$ *are nonempty, then* $A_2^* \subseteq A_1^*$.

(v) *If* $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$ *and* $\lambda > 0$, *then* $(\lambda A)^* = \frac{1}{\lambda} A^*$.

Proof. Part (i) of the theorem is a direct consequence of the definition. Part (ii) can be shown similarly, since

$$\left(\bigcup_{i \in I} A_i\right)^* = \bigcap_{x \in \bigcup_{i \in I} A_i} \{x\}^* = \bigcap_{i \in I} \left(\bigcap_{x \in A_i} \{x\}^*\right) = \bigcap_{i \in I} A_i^*.$$

To prove Part (iii) consider the fact that for any $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$, the set A^* is either $\mathbb{R}^n$ (which is a closed, convex set containing o), or the intersection of closed half spaces containing o. Since closed half spaces are convex sets, and the intersection of closed, convex sets containing o is a closed, convex set containing o, Part (iii) follows. Part (iv) is a consequence of Part (i). Finally, if $A \subseteq \mathbb{R}^n$, $A \neq \emptyset$ and $\lambda > 0$, then

$$\begin{aligned}(\lambda A)^* &= \{y \in \mathbb{R}^n | \langle \lambda x, y\rangle \leq 1 \text{ for every } x \in A\}\\ &= \{y \in \mathbb{R}^n | \langle x, \lambda y\rangle \leq 1 \text{ for every } x \in A\}\\ &= \left\{\frac{1}{\lambda} z \in \mathbb{R}^n \middle| \langle x, z\rangle \leq 1 \text{ for every } x \in A\right\}\\ &= \frac{1}{\lambda}\{z \in \mathbb{R}^n | \langle x, z\rangle \leq 1 \text{ for every } x \in A\} = \frac{1}{\lambda} A^*.\end{aligned}$$

This proves Part (v). □

The next two statements investigate the polars of special classes of sets.

Proposition 11.4. *Let $K \subset \mathbb{R}^n$ be a compact, convex set containing o in its interior. Then K^* is a compact, convex set containing o in its interior.*

Proof. By Part (iii) of Theorem 11.3, K^* is a closed, convex set containing o. We show that K^* is bounded and it contains o in its interior. According to our conditions, there are constants $0 < r < R$ such that $B_r(o) \subseteq K \subseteq B_R(o)$. From this, by Part (iv) of Theorem 11.3 it follows that

$$B_{\frac{1}{R}}(o) = B_R(o)^* \subseteq K^* \subseteq B_r(o)^* = B_{\frac{1}{r}}(o),$$

which yields the statement (Figure 57). □

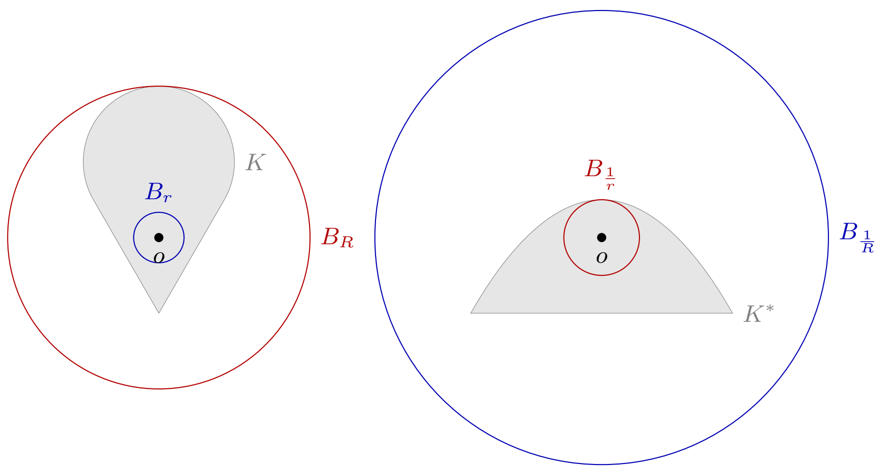


Figure 57: If a closed set K contains o in its interior, then K^* is compact; if K is compact, then $o \in \operatorname{int}(K^*)$.

Proposition 11.5. *Let $K \subseteq \mathbb{R}^n$, $K \neq \emptyset$. Then $(K^*)^* = K$ holds if and only if K is closed, convex, and $o \in K$.*

Proof. If $(K^*)^* = K$, then by Part (iii) of Theorem 11.3, K is closed, convex and $o \in K$. We assume that K is closed, convex and $o \in K$, and show that $(K^*)^* = K$. By the definition of polar, for every $x \in K$ and $y \in K^*$, we have $\langle x, y \rangle \leq 1$, and thus, $K \subseteq (K^*)^*$. Now, let $x \notin K$ be arbitrary. Since K is closed and convex, by Theorem 6.14 there is a hyperplane H that strictly separates x and K (Figure 58). Let H^+ denote the closed half space bounded by H and containing $o \in K$. By the example in the beginning of the lecture, the half space H^+ is the polar of the set $\{y\}$, where the distance of H from o is $\frac{1}{\|y\|}$, and y is an outer normal of H^+. But then $x \notin \{y\}^*$ yields $\langle x, y \rangle > 1$, and $K \subset \{y\}^*$ yields $\langle z, y \rangle \leq 1$ for every $z \in K$. Thus, in this case $y \in K^*$, implying $x \notin (K^*)^*$. This yields $(K^*)^* \subseteq K$, which implies the assertion. $\square$

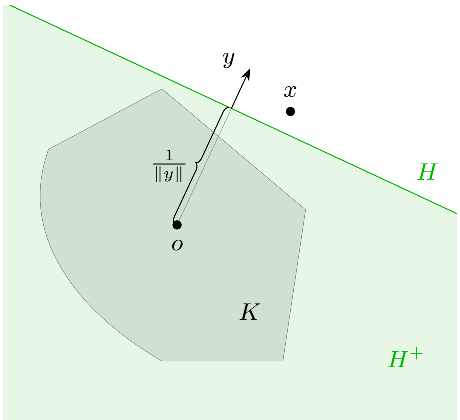


Figure 58: Since $x \in K$ and K is closed and convex, there exists a hyperplane H that strictly separates x and K. The outer normal y of the closed half space H^+ bounded by H and containing $o \in K$ is chosen such that $\text{dist}(o, H) = \frac{1}{\|y\|}$. It follows that $y \in K^*$ and therefore $(K^*)^* \subseteq \{y\}^* = H^+$ does not contain x.

The main result of this lecture is as follows.

Theorem 11.6. *Let $K \subset \mathbb{R}^n$ be a compact, convex set containing o in its interior. To any proper face F of K assign the set*

$$F^\circ = \{y \in K^* | \langle x, y \rangle = 1 \text{ for every } x \in F\}.$$

Then F° is a proper face of K^*, and the map $F \mapsto F^\circ$ is a bijection between the proper faces of K and K^* that reverses containment relation (Figure 59).

Proof. Let $H = \{y \in \mathbb{R}^n | \langle v_0, y \rangle = 1\}$ be an arbitrary supporting hyperplane of K satisfying $F = H \cap K$. Since $\langle v_0, y \rangle \leq 1$ for every $y \in K$ and $\langle v_0, y \rangle = 1$ for every $y \in F$, we have $v_0 \in F^\circ$. Thus, $F^\circ \neq \emptyset$. Now, let $x_0 \in \text{relint}(F)$ and $H' = \{y \in \mathbb{R}^n | \langle y, x_0 \rangle = 1\}$ (Figure 60). By the definition of polar set and $v_0 \in H'$, we have that H' is a supporting hyperplane of K^*, implying that $F' = K^* \cap H'$ is a proper face of K^*. We show that $F' = F^\circ$.

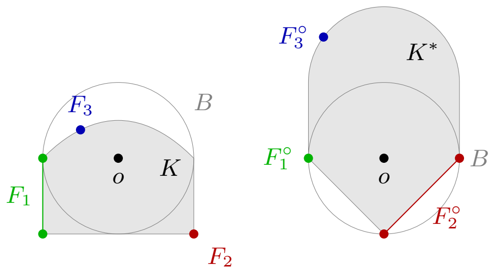


Figure 59: The map $F \mapsto F^\circ$ is a bijection between the faces of K and K^*.

By the definition of F°, $F^\circ \subset H'$ holds, and thus, $F^\circ \subseteq F'$. Now, let $y_0 \in K^* \setminus F^\circ$. Then, there is some $z \in F$ such that $\langle z, y_0 \rangle < 1$. As $x_0 \in \mathrm{relint}(F)$, there is a segment $[z, w] \subseteq F$ such that $x_0 \in [z, w]$ and $x_0 \neq w$. Then x_0 can be written in the form $x_0 = tz + (1-t)w$ for some $t \in (0, 1]$. But $w \in F$ and $y_0 \in K^*$ imply $\langle w, y_0 \rangle \leq 1$, from which

$$\langle x_0, y_0 \rangle = t\langle z, y_0 \rangle + (1-t)\langle w, y_0 \rangle < 1,$$

that is, $y_0 \notin F'$. Thus, we have shown that $F^\circ = F'$ yielding, in particular, that $F \mapsto F^\circ$ is a face of K^*.

Now we prove that for any proper face F, we have $(F^\circ)^\circ = F$, which will imply that the map $F \mapsto F^\circ$ is injective. But since $(K^*)^* = K$; that is, applying this property for K^* we obtain that the map is bijective. By definition,

$$(F^\circ)^\circ = \{y \in (K^*)^* = K | \langle x, y \rangle = 1 \text{ for every } x \in F^\circ\}.$$

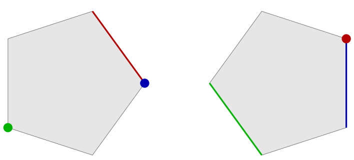
Thus, $F \subseteq (F^\circ)^\circ$. Let us consider the supporting hyperplane $H = \{y \in \mathbb{R}^n | \langle v_0, y \rangle = 1\}$ mentioned in the beginning of the proof. For this hyperplane $H \cap K = F$ is satisfied. During the proof we have shown that $v_0 \in F^\circ$. Hence, if $y \in (F^\circ)^\circ$, then $\langle y, v_0 \rangle = 1$, but by the condition $H \cap K = F$ we have $y \in F$; that is, $(F^\circ)^\circ \subseteq F$.

We need that the map $F \mapsto F^\circ$ reverses the containment relation. But this property is a straightforward consequence of the definition of F°. □

Definition 11.7. Let $P, Q \subset \mathbb{R}^n$ be n-dimensional convex polytopes. We say that Q is a *dual* of P, if there is a bijection between the proper faces of Q and P that reverses containment.

Example 11.8.

(i) Convex polygons are their own duals.



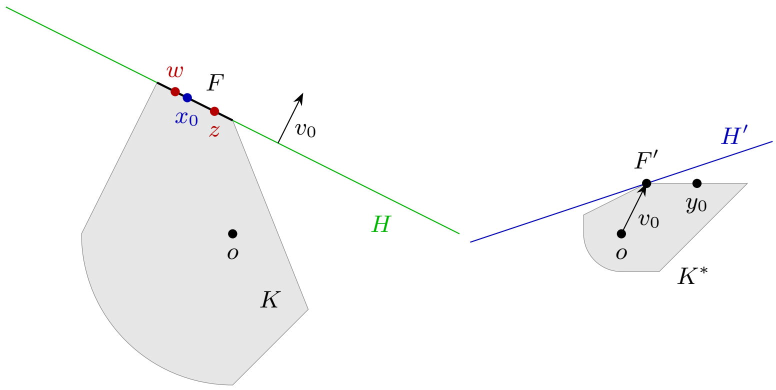
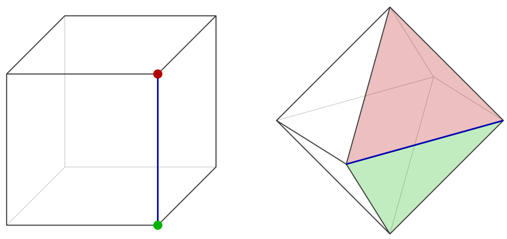
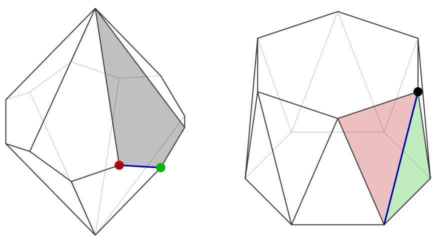


Figure 60: Let F be a face of K and H a supporting hyperplane with outer normal v_0 such that $F = H \cap K$ and $\mathrm{dist}(o, H) = 1/\|v_0\|$. If $x_0 \in \mathrm{relint}(F)$, then $H' = \{y \in \mathbb{R}^n | \langle y, x_0 \rangle = 1\}$ is a supporting hyperplane of K^* containing F°, which in turn contains v_0, therefore $F^\circ \subseteq F' = H' \cap K^*$. If $y_0 \in K^* \setminus F^\circ$, then $\langle z, y_0 \rangle < 1$ for some $z \in F$. Writing x_0 as a convex combination of z and $w \neq x_0$, $w \in F$, we conclude $y_0 \notin F'$, therefore $F' = F^\circ$.

(ii) The octahedron is a dual of the cube.



(iii) The hexagonal antiprism is a dual of the hexagonal trapezohedron.



Problem 11.9. Find dual pairs of polytopes P, Q.

We remark that extending the above map to $\emptyset$ and the polytope itself, the duality of P and Q corresponds to the fact that the face lattices of P and Q are duals (cf. Definition 10.5).

Proposition 11.10. *Let $P \subseteq \mathbb{R}^n$ be an arbitrary convex polytope. Then P has a dual polytope.*

Proof. Since translation and the dimension of the ambient space do not influence the existence of a dual polytope, we may assume that P is n-dimensional, and it contains o in its interior. But then P^* is a dual of P. □

The following statement, which we present without proof, is often used in convex geometry. Before reading it, it is worth recalling that every compact set is Lebesgue measurable, and hence, it has a volume.

Proposition 11.11. *Let K be a compact, convex set containing o in its interior, and let $L : \mathbb{R}^n \to \mathbb{R}^n$ be a nondegenerate linear transformation. Then $V(L(K))V(L(K)^*)$ is independent of the choice of L, where the symbol $V(\cdot)$ denotes the n-dimensional volume.*

Definition 11.12. If $K \subseteq \mathbb{R}^n$ is a compact, convex set containing o in its interior, then the quantity $V(K)V(K^*)$ is called the *volume product* or *Mahler volume* of K.

Theorem 11.13 (Blaschke–Santaló)**.** *For any compact, convex set K with $K = -K$ and $o \in \operatorname{int} K$, we have*

$$V(K)V(K^*) \leq \kappa_n^2 = \frac{\pi^n}{\Gamma\left(\frac{n}{2}+1\right)^2},$$

where κ_n denotes the volume of the n-dimensional unit ball.

The next conjecture is one of the most fundamental conjectures in convex geometry.

Conjecture 11.14 (Mahler)**.** For any compact, convex set K with $K = -K$ and $o \in \operatorname{int} K$, we have

$$V(K)V(K^*) \geq V(C)V(C^*) = \frac{4^n}{n!},$$

where C is a cube centered at o.

It is known that there is an absolute constant $c > 0$ such that $V(K)V(K^*) \geq \frac{c^n}{n!}$ holds for any compact, convex set K with $K = -K$ and $o \in \operatorname{int} K$.

12 Introduction to Hausdorff distance

Our next topic is Hausdorff distance. Let us recall the concepts of Minkowski sum and support function (Definitions 1.1 and 5.4).

If $A, B \subseteq \mathbb{R}^n$ are nonempty sets, then their Minkowski sum is

$$A + B = \{a + b | a \in A, b \in B\}.$$

We have seen that if A, B are compact, convex sets, then $A + B$ is also compact and convex. We have defined the support function of a bounded set K as $h : \mathbb{R}^n \to \mathbb{R}$, $h_K(x) = \sup\{\langle x, y\rangle | y \in K\}$, and we have shown that if $o \in K$, then h_K is convex.

In the lecture we denote the family of compact, convex nonempty sets in $\mathbb{R}^n$ by $\mathcal{K}_n$. The main definition discussed in the lecture is the following.

Definition 12.1. Let $K, L \in \mathcal{K}_n$ be compact sets. Then the *Hausdorff distance* of K and L is (see Figure 61)

$$d_H(K, L) = \inf \{r \geq 0 | K \subseteq L + B_r(o) \text{ and } L \subseteq K + B_r(o)\} .$$

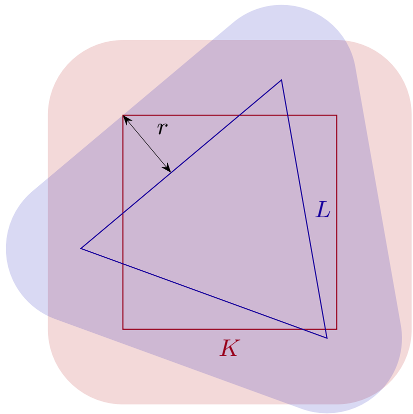


Figure 61: The Hausdorff distance of K and L is the smallest r such that both $K \subseteq L + B_r(o)$ and $L \subseteq K + B_r(o)$.

We remark that the above definition can be extended for bounded sets in general.

Proposition 12.2. *For any $K, L \in \mathcal{K}_n$, we have*

$$d_H(K, L) = \sup \{|h_K(x) - h_L(x)| | x \in \mathbb{R}^n, \|x\| = 1\} .$$

Proof. By Proposition 5.6 we have $K \subseteq L + B_r(o)$ iff for all $x \in \mathbb{R}^n$ we have $h_K(x) \leq h_{L+B_r(o)}(x)$. Using Proposition 5.8 and that $h_{B_r(o)}(x) = r \|x\|$, this is equivalent to $h_K(x) \leq h_L(x) + r \|x\|$ for all x. By property (ii) from Theorem 5.7, this inequality is satisfied iff $h_K(x) \leq h_L(x) + r$ for all unit vectors x. Similarly, $h_L(x) \leq h_{K+B_r(o)}(x)$ iff $h_L(x) \leq h_K(x) + r$ for all unit vectors x. Therefore $|h_L(x) - h_K(x)| \leq r$ is satisfied for all unit vectors x precisely when both $K \subseteq L + B_r(o)$ and $L \subseteq K + B_r(o)$ hold. Now, the statement readily follows by rephrasing the containment relations in the definition of Hausdorff distance. □

Proposition 12.3. *If $K, L, M \in \mathcal{K}_n$, then*

(i) $d_H(K, L) \geq 0$, *with equality if and only if* $K = L$.

(ii) $d_H(K, L) = d_H(L, K)$.

(iii) $d_H(K, L) + d_H(L, M) \geq d_H(K, M)$.

Proof. The inequality $d_H(K, L) \geq 0$ and the equality $d_H(K, K) = 0$ follows from the definition. On the other hand, if $d_H(K, L) = 0$, then $K \subseteq L$ and $L \subseteq K$, implying $K = L$. The definition does not disinguish the order of K and L, and thus, $d_H(K, L) =$

$d_H(L,K)$. Finally, if $K \subseteq L + B_{r_1}(o)$ and $L \subseteq M + B_{r_2}(o)$, then $B_{r_1}(o) + B_{r_2}(o) = B_{r_1+r_2}(o)$ yields $K \subseteq M + B_{r_1+r_2}(o)$, and $M \subseteq L + B_{r_2}(o)$ and $L \subseteq K + B_{r_1}(o)$ implies similarly that $M \subseteq K + B_{r_1+r_2}(o)$. From this we obtain the triangle inequality $d_H(K,L) + d_H(L,M) \geq d_H(K,M)$. □

Corollary 12.4. *The family $\mathcal{K}_n$, equipped with Hausdorff distance, is a metric space.*

Let us recall that a metric space is called a *complete metric space* if every Cauchy sequence in the space is convergent. This property is investigated in the next theorem.

Theorem 12.5. *The family $\mathcal{K}_n$, equipped with Hausdorff distance, is a complete metric space.*

Proof. Let $K_i \in \mathcal{K}_n$, $i = 1, 2, \ldots$ be a Cauchy sequence of nonempty, compact, convex sets; i.e. assume that for every $\varepsilon > 0$ there is some $m_0 \in \mathbb{Z}^+$ such that if $m_1, m_2 > m_0$, then $d_H(K_{m_1}, K_{m_2}) < \varepsilon$. We show that then there is some $K \in \mathcal{K}_n$ such that $K_m \to K$ with respect to the topology induced by Hausdorff distance.

For every positive integer i, let $B_i = \mathrm{cl}(K_i \cup K_{i+1} \cup \ldots)$. By the properties of Cauchy sequences, B_i is a nonempty, bounded and closed set in $\mathbb{R}^n$, implying that it is compact, and $B_{i+1} \subseteq B_i$ for every i. Let $B = \cap_{i=1}^{\infty} B_i$. Since the intersection of arbitrarily many closed sets is closed, B is compact. We show that B is not empty. Indeed, if $B = \emptyset$, then the complements of the sets B_i with respect to the compact set B_1 form an open cover of B_1. But then we can choose a finite open subcover of B_1, i.e. there are finitely many B_is whose intersection is $\emptyset$, from which, as the sets are nested, it follows that $B_i = \emptyset$ for some value of i, which contradicts the definition of B_i. We have obtained that B is a nonempty, compact set.

Let $\varepsilon > 0$ be arbitrary. We show that there is an index $m \in \mathbb{Z}^+$ such that for every $i > m$, we have $B_i \subseteq \mathrm{int}(B + B_\varepsilon(o))$. By contradiction, suppose that it is not true. Then there is a sequence i_j of indices such that for every value of j, $B_{i_j} \not\subseteq \mathrm{int}(B + B_\varepsilon(o))$. Let $C_{i_j} = B_{i_j} \setminus \mathrm{int}(B + B_\varepsilon(o))$. By our conditions, the sets C_{i_j} are nonempty, nested, compact sets, which implies, as in the previous paragraph, $C = \cap_{i=1}^{\infty} C_{i_j}$ is a nonempty, compact set. But as the sets B_i are nested, $C \subseteq B_{i_j}$ for every value of j, implying that $C \subseteq B_i$ for every value of i. On the other hand, by their constructions, C and B are disjoint, which is a contradiction. Thus, for a suitable $m \in \mathbb{Z}^+$, $B_i \subseteq \mathrm{int}(B + B_\varepsilon(o))$ for all $i > m$. But from this it follows that $K_i \subseteq B + B_\varepsilon(o)$ for all $i > m$.

Since $\{K_i\}$ is a Cauchy sequence, there is an index k such that $d_H(K_i, K_j) < \varepsilon$ if $i, j > k$. Thus, if $i > k$ is arbitrary, then $\bigcup_{j=i}^{\infty} K_i \subseteq K_i + B_\varepsilon(o)$, implying $B \subseteq B_i \subseteq K_i + B_\varepsilon(o)$. This yields that if $i > \max\{k, m\}$, then $d_H(B, K_i) \leq \varepsilon$, and thus, the limit set of $\{K_i\}$ is B.

We need to show that B is convex. Let $p, q \in B$ be arbitrary, and assume that for some $t \in (0,1)$, $x = tp + (1-t)q \notin B$. Then, by the compactness of B, there is a value $\delta > 0$ such that $B_\delta(x) \cap B = \emptyset$. Since the limit set of $\{K_i\}$ is B, there is an index i such that $K_i \subseteq B + B_{\delta/2}(o)$ and some points $p', q' \in K_i$ such that $\|p - p'\|, \|q - q'\| \leq \frac{\delta}{2}$. Let $x' = tp' + (1-t)q' \in K_i$, which, by the triangle inequality, implies that $\|x - x'\| \leq t\|p - p'\| + (1-t)\|q - q'\| \leq \frac{\delta}{2}$, and thus, $x \in B_{\delta/2}(x')$. But from this we obtain

$x \in K_i + B_{\delta/2}(o) \subseteq B + B_\delta(o)$, or in other words, $B_\delta(x) \cap B \neq \emptyset$, which is in contradiction with the choice of δ. □

Definition 12.6. Let $\mathcal{F}$ be a nonempty family of nonempty sets in $\mathbb{R}^n$. If there is some $r > 0$ such that $F \subseteq B_r(o)$ for every $F \in \mathcal{F}$, then we say that $\mathcal{F}$ is *uniformly bounded.*

The next theorem is a generalization of the Bolzano–Weierstrass theorem for bounded sequences.

Theorem 12.7 (Blaschke's selection theorem)**.** *Let $\mathcal{F} \subseteq \mathcal{K}_n$ be a uniformly bounded, infinite family. Then $\mathcal{F}$ contains a sequence converging to an element of $\mathcal{K}_n$.*

Proof. We show that $\mathcal{F}$ contains a Cauchy sequence. Let C be a cube in $\mathbb{R}^n$ that contains all elements of $\mathcal{F}$, and let the edge length of C be r. Let i be a positive integer, and dissect C with hyperplanes parallel to its facets into smaller (closed) cubes of edge length $\frac{r}{2^i}$. To any element K of $\mathcal{F}$, assign the union of the small cubes that intersect K. We call this set the ith minimal cover (Figure 62).

Since there are only finitely many possible first minimal covers, there is a union F_1 of small cubes which is the first minimal cover of infinitely many elements of $\mathcal{F}$. Let $\mathcal{F}_1 \subset \mathcal{F}$ be the subset of $\mathcal{F}$ whose first minimal cover is F_1. As $|\mathcal{F}_1| = \infty$ and there are only finitely many possible second minimal covers, there is a union F_2 of small cubes that is the second minimal cover of infinitely many elements of $\mathcal{F}_1$. Continuing this process, we obtained a sequence of nested subfamilies $\mathcal{F} \supseteq \mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \ldots \supseteq \mathcal{F}_i \supseteq \ldots$ whith the property that every element of $\mathcal{F}_i$ has the same ith minimal cover F_i.

Let $K_i \in \mathcal{F}_i$, and consider the sequence $\{K_i\}$. According to the construction, for any $K_i \in \mathcal{F}_i$, $K_j \in \mathcal{F}_j$, $i < j$, the ith minimal cover of K_i and K_j i coincides. Since the diameters of the cubes forming an ith minimal cover is $\frac{r\sqrt{n}}{2^i}$, therefore then $d_H(K_i, K_j) \leq \frac{r\sqrt{n}}{2^i}$ (Figure 63). But this implies that $\{K_i\}$ is a Cauchy sequence, and thus, by the previous theorem, it is convergent. □

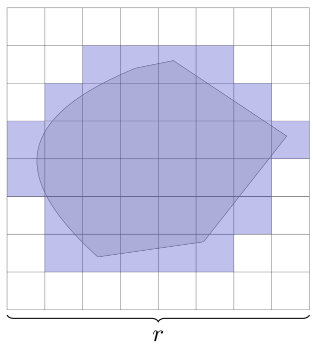


Figure 62: A compact convex set in the plane and its third minimal cover with respect to a square of edge length r.

According to the next theorem, the family of convex polytopes is an everywhere dense subfamily in $\mathcal{K}_n$.

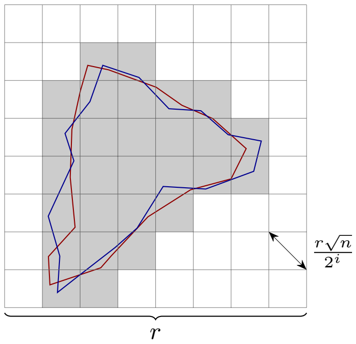


Figure 63: If two sets have the same ith minimal cover, then their Hausdorff distance is at most $\frac{r\sqrt{n}}{2^i}$, the diameter of an n-dimensional cube of edge length $\frac{r}{2^i}$.

Theorem 12.8. *Let $K \in \mathcal{K}_n$ be arbitrary. Then there is a sequence of convex polytopes $\{P_k\}$ that converges to K with respect to Hausdorff distance.*

Proof. Without loss of generality, assume that $\dim(K) = n$. To prove the statement, it is sufficient to show that for every $\varepsilon > 0$ there is some convex polytope P satisfying $P \subseteq K \subseteq P + B_\varepsilon(o)$, since choosing a polytope P_k for every positive integer k with the property that $P_k \subseteq K \subseteq P_k + B_{1/k}(o)$, the sequence $\{P_k\}$ satisfies the required conditions.

Since K is compact, there are points $x_1, \ldots, x_m \in K$ such that the open balls $\operatorname{int} B_\varepsilon(x_i)$ cover K. Let $P = \operatorname{conv}\{x_1, \ldots, x_m\}$ (Figure 64). Then, clearly $P \subseteq K$. But $K \subseteq \bigcup_{i=1}^m \operatorname{int}(B_\varepsilon)(x_i) = \{x_1, \ldots, x_m\} + \operatorname{int} B_\varepsilon(o) \subseteq P + B_\varepsilon(o)$, from which the assertion follows. □

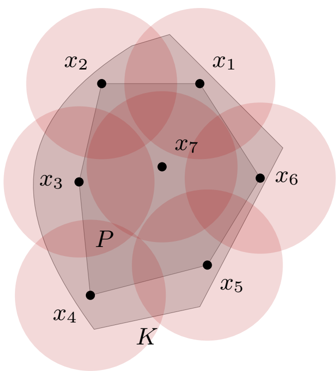


Figure 64: Since K is compact, for every $\epsilon > 0$ there exist finitely many points $x_1, \ldots, x_m \in K$ such that the union of the ϵ-balls centered at these points contains K. As K is convex, it follows that the convex polytope $P = \operatorname{conv}\{x_1, \ldots, x_m\}$ has a Hausdorff distance of at most ϵ to K.