CONVEXITY

© Alexander Barvinok

2001

Abstract. These are the lecture notes for the graduate course Math 669, which I taught at the University of Michigan in the Winter term of 2001.

Typeset by $\mathcal{AMS}$-TEX

Contents

1. Convex sets. Some interesting examples and problems. We introduce the notions of Euclidean space, vector space, convex combination, convex hull and convex set. Some of the examples and problems preview topics of the course. 7

2. Properties of the convex hull. Carathéodory's Theorem. We establish some elementary properties of convex hulls. Then we prove that every point in the convex hull of a set in $\mathbb{R}^d$ is a convex combination of some $d+1$ points of the set (Carathéodory's Theorem). As a corollary, we prove that the convex hull of a compact in $\mathbb{R}^d$ is compact. Problems include Gauss' Theorem relating the roots of a complex polynomial and the roots of its derivative as well as some interesting extensions of Carathéodory's Theorem. 11

3. Theorems of Radon and Helly. The main result of the section is Helly's Theorem, which asserts that a family of convex sets in $\mathbb{R}^d$ has a non-empty intersection provided the intersection of every $d+1$ members of the family is non-empty (if the family is infinite, compactness may be required). Problems contain various extensions of the theorem from quite straightforward to rather involved (intersection with a subspace, fractional Helly's Theorem, "Piercing" Theorem, volume of the intersection, etc.) 16

4. Some applications of Helly's Theorem. We discuss when two sets of points in Euclidean space can be separated by a hyperplane and why any mass distribution in $\mathbb{R}^d$ has a "center" such that every hyperplane through that center cuts at least $1/(d+1)$ of the total mass. More applications ("Ham Sandwich Theorem", Krasnoselsky's Theorem, Jung's Theorem) are in the exercises. 19

5. An application to approximation. We show that a real-valued function can be approximated within a given error by a linear combination of some $d+1$ given functions uniformly on a set X if it can be so approximated at any given $d+2$ points of X (if X is infinite, some mild conditions may apply). 22

6. Euler characteristic. We define the Euler characteristic and prove its existence as the (necessarily unique) linear functional $\mu : \mathcal{C}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ on the vector space $\mathcal{C}(\mathbb{R}^d)$ spanned by the indicator functions $[A]$ of closed convex subsets $A \subset \mathbb{R}^d$ such that $\mu([A]) = 1$ for every non-empty closed convex $A \subset \mathbb{R}^d$. 25

7. Application: convex sets and linear transformations. We prove that linear dependences among indicator functions of convex compact sets in $\mathbb{R}^d$ are preserved under linear transformations. As an application, we construct intrinsic volumes ω_k: $0 \leq k \leq d$ of convex compact sets in $\mathbb{R}^d$, which are additive, rotation invariant and homogeneous of degree k functions. As an exercise, we describe a lifting of the Minkowski addition of convex compact sets in $\mathbb{R}^d$ to a bilinear operation on the vector space spanned by indicator functions of such sets. 30

8. The Isolation Theorem. We define affine subspaces, affine hyperplanes, affine combinations and affine hulls in an arbitrary real vector space V. A convex set $A \subset V$ is called algebraically open if every straight line intersects A by an open interval (possibly empty). We prove that for any convex algebraically open set $A \subset V$ and a point $u \notin A$ there is an affine hyperplane H containing u and having set A "on one side". This result, called the Isolation Theorem, will

be one of our main technical tools. An exercise shows that the condition of algebraic openness can not be removed (in the infinite dimensional case). 34

9. Convex sets in Euclidean space. We introduce the notion of a face of a convex set and prove a number of results regarding the structure, especially the boundary structure, of convex sets in Euclidean space. The main difference with the infinite dimensional situation is that every convex set in Euclidean space is full-dimensional when considered in its affine hull. One of the main results of this section is the existence of a support hyperplane at every point of the boundary of a convex set. 39

10. Extreme points. The Krein-Milman Theorem in Euclidean space. We define extreme points of convex sets and prove that a convex compact set in Euclidean space is the convex hull of the set of its extreme points. 42

11. Extreme points of polyhedra. A polyhedron $P \subset \mathbb{R}^d$ is a set of solutions to a system of finitely many linear inequalities. We prove that at every extreme point of P at least d of those inequalities hold with equality and describe the appearance of "unrealistic" solutions in linear programming problems as a consequence of that. Also, we prove that a bounded polyhedron is a polytope (the convex hull of finitely many points). 44

12. The Birkhoff polytope. We consider the set (Birkhoff polytope) of all $n \times n$ doubly stochastic matrices (that is, non-negative matrices with the row and column sums equal to 1) and prove that the extreme points of the Birkhoff polytope are the $n \times n$ permutation matrices (the Birkhoff - von Neumann Theorem). 46

13. The permutation polytope and the Schur-Horn Theorem. We define the permutation polytope $P(x) \subset \mathbb{R}^n$ as the convex hull of all points obtained from $x \in \mathbb{R}^n$ by a permutation of the coordinates. We prove that a diagonal of an $n \times n$ symmetric matrix A with the prescribed eigenvalues $\lambda_1, \ldots, \lambda_n$ must lie in $P(l)$ for $l = (\lambda_1, \ldots, \lambda_n)$ (Schur's Theorem). 48

14. The transportation polyhedron. Given a graph (network) with real numbers (interpreted as demands when positive and supplies when negative) at the vertices, we define the transportation polyhedron as the set of all feasible (satisfying the demand/supply and non-negativity constraints) flows. We show that extreme points of the polyhedron correspond to flows without circulation. As a corollary, we prove that the extreme points are integral provided the demands/supplies are integral. 50

15. Convex cones. We define conic combinations, conic hulls and convex cones. We define extreme rays and prove a conic version of the Krein-Milman Theorem. 53

16. The moment curve and the moment cone. The curve $g(\tau) = (1, \tau, \tau^2, \ldots, \tau^d)$, $\alpha \leq \tau \leq \beta$ is called the moment curve in $\mathbb{R}^{d+1}$ and its conic hull is called the moment cone. The main result of this section is that every point u from the moment cone in $\mathbb{R}^{d+1}$ can be written as a non-negative linear combination of $(d+1)/2$ when d is odd and $(d+2)/2$ when d is even points on the moment curve. Problems explore how unique the situation of the moment cone is. 56

17. An application: "double precision" formulas for numerical integration. Let $\rho(\tau)$ be a continuous non-negative density on the interval $[\alpha, \beta]$. We prove that for every $m \geq 1$ there are m nodes $\tau_1^*, \ldots, \tau_m^*$ in $[\alpha, \beta]$ and m non-negative weights $\lambda_1, \ldots, \lambda_m$ such that $\int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \sum_{i=1}^m \lambda_i f(\tau_i^*)$ for every polynomial f of degree at most $2m-1$. Problems

describe such weights and nodes for low degree polynomials, show the uniqueness for strictly positive densities ρ and explore some generalizations. 58

18. The cone of non-negative polynomials. We describe the extreme rays of the convex cone of all polynomials $p(\tau)$ of degree d that are non-negative on the (bounded) interval $[\alpha, \beta]$ and obtain representations of such polynomials as sums of squares of polynomials, possibly multiplied by $(\tau - \alpha)$, $(\beta - \tau)$ or $(\tau - \alpha)(\beta - \tau)$. Problems treat the cases of unbounded intervals and sharpen the representation theorems. 62

19. The cone of positive semidefinite matrices. The main result is that the faces of the cone of positive semidefinite $n \times n$ matrices are parameterized by the subspaces of $\mathbb{R}^n$: if $L \subset \mathbb{R}^n$ is a subspace of codimension r, then the set of positive semidefinite matrices whose null space contains L is a face of dimension $r(r+1)/2$. 66

20. Linear equations in positive semidefinite matrices. Let r and $k < (r+1)(r+2)/2$ be non-negative integers. We prove that if a system of k linear equations in $n \times n$ symmetric matrices has a positive semidefinite solution, then there is a positive semidefinite solution whose rank does not exceed r. We prove that if $r > 0$, $n \geq r+2$ and the set of all positive semidefinite solutions is non-empty and bounded, there is a positive semidefinite solution of rank at most r even when $k = (r+1)(r+2)/2$. Problems show that the bounds, generally speaking, can not be strengthened. 70

21. Applications: quadratic convexity theorems. We prove that the image of the unit sphere in $\mathbb{R}^n$ under a quadratic map $\mathbb{R}^n \longrightarrow \mathbb{R}^k$ is a convex set in $\mathbb{R}^k$ for $k = 2$ provided $n \geq 3$, whereas for $k > 2$ every point from the convex hull of the image is a convex combination of a small number (about $O(\sqrt{k})$) points from the image. Problems contain various extensions, in particular, to complex Hermitian maps, ramifications and connections to formulas for numerical integration, multiplicities of eigenvalues, etc. 76

22. Applications: problems of graph realizability. Let $G = (V, E)$ be a graph with $n = |V|$ vertices, $k = |E|$ edges and non-negative weights on the edges. We say that G is realizable in $\mathbb{R}^d$ if there is a map $\phi : V \longrightarrow \mathbb{R}^d$ such that the Euclidean distance between $\phi(u)$ and $\phi(v)$ is equal to the prescribed weight whenever $(u, v) \in E$ is an edge of G. The main result is that if $k < (d+1)(d+2)/2$ then G is realizable in some Euclidean space if and only if it is realizable in $\mathbb{R}^d$. Moreover, the same conclusion holds true if $k = (d+1)(d+2)/2$ and G is not a union of a complete graph with $d+2$ vertices and zero or more isolated vertices. Problems explore various special cases and the convexity of the image of the rigidity map, a quadratic map originating in the space of all possible embeddings of a given graph. 81

23. Separation theorems in Euclidean space. We prove that two non-intersecting convex sets in Euclidean space can be separated by a hyperplane. The result holds true in a general vector space if we assume that at least one of the sets is algebraically open. A problem shows that without that assumption the result is generally false. 85

24. Topological vector spaces, convex sets and hyperplanes. We define topological vector spaces and prove that the interior and closure of a convex set is convex and that a hyperplane is either closed or dense. 89

25. Separation theorems in topological vector spaces. We introduce locally convex topological vector spaces. The main result is that a closed convex set A in a locally convex topological

vector space V can be strictly separated from a point $u \notin A$ by a closed hyperplane. 95

26. The Krein-Milman Theorem in topological vector spaces. We prove that a convex compact set in a locally convex topological vector space is the convex hull of the set of its extreme points (the Krein-Milman Theorem). 98

27. Polyhedra in L^∞. We consider a set A of functions $f \in L^\infty[0,1]$ satisfying finitely many integral constraints and such that $0 \le f(\tau) \le 1$ for almost all $\tau \in [0,1]$ (thus A may be considered as an "infinite-dimensional cube".) We prove that if f is an extreme point of A then $f(\tau) \in \{0,1\}$ for almost all τ. 100

28. An application: problems of linear optimal control. We show that in a "typical" solution of a linear optimal control problem with the control function $u \in L^\infty[0,1]$ one has $u(\tau) \in \{0,1\}$ for almost all τ. 103

29. An application: the Lyapunov Convexity Theorem. We prove a special, but sufficiently general, case of the Lyapunov Convexity Theorem, which asserts that the range of a vector non-atomic countably additive measure with values in $\mathbb{R}^m$ is a convex set in $\mathbb{R}^m$. As a corollary, we deduce that the set of values of the vector integral functional $\int_0^1 F\big(x(\tau),\tau\big)\,d\tau$, taken over all trajectories $x(\tau) \in L^\infty[0,1]$ that stay within a prescribed set $S \subset \mathbb{R}$ for almost all τ, is convex. 106

30. The "simplex" of probability measures. We define the set Δ of Borel probability measures on the interval $[0,1]$ as the set of non-negative linear functionals on the space $C[0,1]$ of continuous real-valued functions. We prove that the extreme points of Δ are the delta-measures δ_x such that $\delta_x(f) = f(x)$ for every $f \in C[0,1]$ and argue that Δ can be viewed as an infinite-dimensional version of a simplex. 109

31. Extreme points of an intersection. Applications. Let V be a vector space and let $K \subset V$ be a convex set such that for every straight line $L \subset V$, the intersection $K \cap L$ is a closed bounded interval, possibly empty or a point. Let $A \subset V$ be an affine subspace of finite codimension m. We prove that every extreme point of the intersection $K \cap A$ is a convex combination of at most $m+1$ extreme points of K. As an application, we prove that the extreme probability measures satisfying m linear constraints are convex combinations of $m+1$ delta-measures. 113

32. Polarity in Euclidean space. We define the polar A° of a set $A \subset \mathbb{R}^d$ and prove that $(A^\circ)^\circ = A$ provided A is a convex closed set containing the origin. As a corollary, we deduce that a polytope is a polyhedron. We prove that polarity preserves linear relations among indicator functions of closed convex cones and argue that polarity can be considered as the Fourier transform with the Euler characteristic as a measure. Problems touch on the polarity correspondence in the projective space, self-dual polytopes and other topics. 117

33. An application: recognizing points in the moment cone. We find necessary and sufficient conditions for the existence of a Borel measure μ on the interval $[0,1]$ with the prescribed first d moments. The conditions reduce to checking positive semidefiniteness of a pair of explicitly constructed quadratic forms. 123

34. Polytopes and polarity. We establish an inclusion reversing bijection between the faces of a polytope P and the faces of its polar P° (assuming that 0 is in the interior of P). This

correspondence is further explored in the problems. In particular, the set of faces $F \supset G$ of a d-dimensional polytope P containing a given k-dimensional face G of P has the combinatorial structure of a $(d-k-1)$-dimensional polytope. 127

35. The facial structure of the permutation polytope. We describe the facial structure of the permutation polytope $P(a) \subset \mathbb{R}^n$ assuming that the coordinates of a are distinct. As a corollary, we find a necessary and sufficient condition for a point $b = (\beta_1, \ldots, \beta_n)$ to lie in the permutation polytope $P(a)$: assuming that $\beta_1 \geq \ldots \geq \beta_n$ and $\alpha_1 \geq \ldots \geq \alpha_n$, we have $b \in P(a)$ if and only if $\sum_{i=1}^{k} \beta_i \leq \sum_{i=1}^{k} \alpha_i$ for $k = 1, \ldots, n-1$ and $\sum_{i=1}^{n} \beta_i = \sum_{i=1}^{n} \alpha_i$ (Rado's Theorem). 132

36. The Euler-Poincaré Formula. We define a polytopal complex as a finite collection $\mathcal{P}$ of distinct polytopes in $\mathbb{R}^d$ with the property that the intersection of any two polytopes P_1 and P_2 from $\mathcal{P}$ is a common face of both P_1 and P_2 and that a face of a polytope in $\mathcal{P}$ also belongs to $\mathcal{P}$. Let $|\mathcal{P}|$ be the union of the polytopes in a polytopal complex $\mathcal{P}$. We prove that the Euler characteristic of $|\mathcal{P}|$ is the alternating sum $\sum_{i=0}^{d} (-1)^i f_i(\mathcal{P})$, where $f_i(\mathcal{P})$ is the number of i-dimensional polytopes in $\mathcal{P}$. As an application we deduce the Euler-Poincaré formula relating the numbers of faces of various dimensions of a polytope. Problems include various extensions and reformulations, for example, the combinatorial version of the Gauss-Bonnet formula for 3-dimensional polytopes. 135

37. Polytopes with many faces: cyclic polytopes. Let us choose n distinct points $v_1, \ldots, v_n$ on the curve $(\tau, \tau^2, \ldots, \tau^d)$: $0 \leq \tau \leq 1$ in $\mathbb{R}^d$. The convex hull of $v_1, \ldots, v_n$ is called the cyclic polytope $C(d,n)$. We prove that every subset of at most $d/2$ vertices of $C(d,n)$ form a face of $C(d,n)$. Problems describe the facial structure of the cyclic polytope in some more detail, and, in particular, justify the term "cyclic". 140

38. Simple polytopes. A d-dimensional polytope P is called simple if every vertex of P belongs to precisely d facets. Locally, in the neighborhood of every vertex, a simple polytope looks like the non-negative orthant $\mathbb{R}^d_+$. 141

39. The h-vector of a simple polytope. Dehn-Sommerville equations. Let $P \subset \mathbb{R}^d$ be a d-dimensional simple polytope and let $\ell : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function which is not constant on edges of P. We define the index of a vertex v of P as the number of neighbors u (so that $[u,v]$ is an edge of P) such that $\ell(u) < \ell(v)$ and the k-th h-number $h_k(P)$ as the number of the vertices of P having index k. We express $h_k(P)$ solely in terms of the numbers of faces of P of various dimension, implying, in particular, that $h_k(P)$ does not depend on the choice of ℓ. The Dehn-Sommerville equations state that $h_k(P) = h_{d-k}(P)$. Problems address, in particular, combinatorial interpretation of $h_k(P)$, when P is a cube or permutation polytope. 144

40. The Upper Bound Theorem. We find the largest possible number of the k-dimensional faces that a d-dimensional simple polytope with n facets may have. The same bound holds for not necessarily simple polytopes (sketched in a problem). 147

1. Convex Sets. Some Interesting Examples and Problems

Many (though not all) things that happen in this course, happen in Euclidean space $\mathbb{R}^d$.

(1.1) Euclidean space. The d-dimensional Euclidean space $\mathbb{R}^d$ consists of all d-tuples $x = (\xi_1, \dots, \xi_d)$ of real numbers. A an element of $\mathbb{R}^d$ is called a *vector* or a *point*. We can add points: we say that $z = x + y$ for $x = (\xi_1, \dots, \xi_d)$, $y = (\eta_1, \dots, \eta_d)$ and $z = (\zeta_1, \dots, \zeta_d)$, provided $\zeta_i = \xi_i + \eta_i$ for $i = 1, \dots, d$. We can multiply a point by a real number: if $x = (\xi_1, \dots, \xi_d)$ and α is a real number, then $\alpha x = (\alpha\xi_1, \dots, \alpha\xi_d)$ is a point from $\mathbb{R}^d$. We consider the *scalar product* in $\mathbb{R}^d$: $\langle x, y\rangle = \xi_1\eta_1 + \dots + \xi_d\eta_d$, where $x = (\xi_1, \dots, \xi_d)$ and $y = (\eta_1, \dots, \eta_d)$. We define the norm $\|x\| = \sqrt{\xi_1^2 + \dots + \xi_d^2}$ of a point $x \in \mathbb{R}^d$ and the distance between two points x and y: $\operatorname{dist}(x, y) = \|x - y\|$ for $x, y \in \mathbb{R}^d$.

We introduce the central concept of the course right away.

(1.2) Convex sets, convex combinations, and convex hulls.

Let $\{x_1, \dots, x_m\}$ be a finite set of points from $\mathbb{R}^d$. A point

$$x = \sum_{i=1}^{m} \alpha_i x_i, \qquad \text{where} \quad \sum_{i=1}^{m} \alpha_i = 1 \quad \text{and} \quad \alpha_i \geq 0 \quad \text{for} \quad i = 1, \dots, m$$

is called a *convex combination* of $x_1, \dots, x_m$. Given two distinct points $x, y \in \mathbb{R}^d$, the set

$$[x, y] = \Big\{\alpha x + (1 - \alpha)y : 0 \leq \alpha \leq 1\Big\}$$

of all convex combinations of x and y is called the *interval* with the endpoints x and y. A set $A \subset \mathbb{R}^d$ is called *convex*, provided $[x, y] \subset A$ for any two $x, y \in A$. We agree that the empty set $\emptyset$ is convex. For $A \subset \mathbb{R}^d$, the set of all convex combinations of points from A is called the *convex hull* of A and denoted $\operatorname{conv}(A)$. We will see that $\operatorname{conv}(A)$ is the smallest convex set containing A.

(1.3) Some interesting examples. Sometimes, it is very easy to see whether

the set is convex or not.

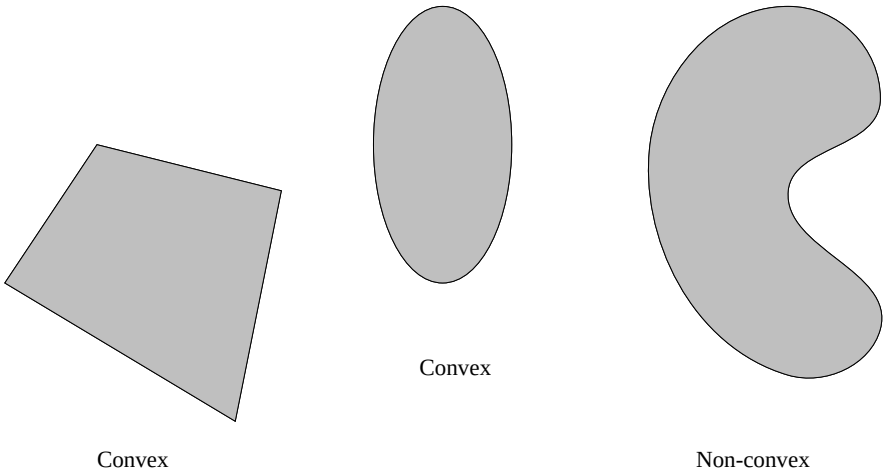


Figure 1

Sometimes, however, this is not so easy to see (cf. Problems 4 and 5 below). Sometimes, a convex set may have a couple of equivalent descriptions and their equivalence is not immediate to see (cf. Problems 6 and 7 below).

PROBLEMS.

1°. Prove that the convex hull of a set is a convex set.

2°. Let $c_1, \ldots, c_n$ be vectors from $\mathbb{R}^d$ and let $\alpha_1, \ldots, \alpha_n$ be numbers. The set

$$A = \Big\{ x \in \mathbb{R}^d : \langle c_i, x \rangle \leq \alpha_i \quad \text{for} \quad i = 1, \ldots, m \Big\}$$

is called a *polyhedron*. Prove that a polyhedron is a convex set.

3 (Gauss' Theorem). Let $f(z)$ be a non-constant polynomial of one complex variable $z \in \mathbb{C}$ and let $z_1, \ldots, z_n$ be the roots of f, so that $f(z_1) = \ldots = f(z_n) = 0$. Suppose that w is a root of f', that is $f'(w) = 0$. Prove that $w \in \mathrm{conv}\{z_1, \ldots, z_n\}$ (we identify $\mathbb{C} = \mathbb{R}^2$).

4* (A version of the Toeplitz-Hausdorff Theorem). Let $q_1, q_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms and let $S^{n-1} = \big\{ x \in \mathbb{R}^n : \|x\| = 1 \big\}$ be the unit sphere. Consider the map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^2$, $T(x) = \big(q_1(x), q_2(x)\big)$. Prove that the image $T(S^{n-1})$ of the sphere is a convex set in $\mathbb{R}^2$, provided $n > 1$.

5* (The Schur-Horn Theorem). For an $n \times n$ real symmetric matrix $A = (a_{ij})$, let $\mathrm{diag}(A) = (a_{11}, \ldots, a_{nn})$ be the diagonal of A, considered as a vector from $\mathbb{R}^n$. Let us fix real numbers $\lambda_1, \ldots, \lambda_n$. Consider the set $X \subset \mathbb{R}^n$ of all diagonals of $n \times n$ real symmetric matrices with the eigenvalues $\lambda_1, \ldots, \lambda_n$. Prove that X is a convex set. Furthermore, let $l = (\lambda_1, \ldots, \lambda_n)$ be the vector of eigenvalues, so $l \in \mathbb{R}^n$. For a permutation σ of the set $\{1, \ldots, n\}$, let $l^\sigma = (\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)})$, $l^\sigma \in \mathbb{R}^n$. Then

$$X = \mathrm{conv}\Big\{ l^\sigma : \sigma \quad \text{ranges over all } n! \text{ permutations of the set } \{1, \ldots, n\} \Big\}.$$

6 (The Birkhoff - von Neumann Theorem). For a permutation σ of the set $\{1, \ldots, n\}$, let us define the $n \times n$ permutation matrix $X^\sigma = (x_{ij}^\sigma)$ as

$$x_{ij}^\sigma = \begin{cases} 1 & \text{if } \sigma(i) = j, \\ 0 & \text{if } \sigma(i) \neq j. \end{cases}$$

Prove that the convex hull of all $n!$ permutation matrices X^σ is the set of all $n \times n$ doubly stochastic matrices X:

$$X = \left\{ (x_{ij}) : \sum_{i=1}^{n} x_{ij} = 1 \quad \text{for all } j \right.$$
$$\sum_{j=1}^{n} x_{ij} = 1 \quad \text{for all } i$$
$$\left. x_{ij} \geq 0 \quad \text{for all } \ i, j \right\}.$$

We consider an $n \times n$ matrix as a point in $\mathbb{R}^{n^2}$.

7. Let us fix an even number $n = 2m$ and let us interpret $\mathbb{R}^{n+1} = \{(\alpha_0, \ldots, \alpha_n)\}$ as the space of all polynomials $p(\tau) = \alpha_0 + \alpha_1 \tau + \ldots + \alpha_n \tau^n$ of degree at most n in one variable τ. Let

$$K = \left\{ p \in \mathbb{R}^{n+1} : p(\tau) \geq 0 \quad \text{for all } \ \tau \in \mathbb{R} \right\}$$

be the set of all non-negative polynomials. Prove that K is a convex set and that K is the set of all polynomials that are representable as sums of squares of polynomials of degree at most m:

$$K = \left\{ \sum_{i=1}^{k} q_i^2, \quad \text{where} \quad \deg q_i \leq m \right\}.$$

Sometimes, it is useful to consider convex sets in a more general setting.

(1.4) Convex sets in vector spaces. We recall that a set V with the operations "+" (addition): $V \times V \longrightarrow V$ and "$\cdot$" (multiplication) $\mathbb{R} \times V \longrightarrow V$ is called a *vector space* provided the following axioms are satisfied: $u + v = v + u$ for any two $u, v \in V$, $u + (v + w) = (u + v) + w$ for any three $u, v, w \in V$, there exists a zero vector $\mathbf{0} \in V$ such that $v + \mathbf{0} = v$ for each $v \in V$, for each $v \in V$ there is $-v \in V$ such that $v + (-v) = \mathbf{0}$, $(\alpha\beta)v = \alpha(\beta v)$ for any $v \in V$, any $\alpha, \beta \in \mathbb{R}$, $1v = v$ for any $v \in V$, $(\alpha + \beta)v = \alpha v + \beta v$ for any $v \in V$, $\alpha, \beta \in V$, and $\alpha(v + u) = \alpha v + \alpha u$ for any $\alpha \in \mathbb{R}$ and any $u, v \in V$. We often say "points" instead of "vectors".

A set $A \subset V$ is called *convex*, provided for each $x, y \in A$ the interval

$$[x, y] = \{\alpha x + (1 - \alpha)y : 0 \leq \alpha \leq 1\}$$

is contained in A. Again, we agree that the empty set is convex.

PROBLEM

1°. Let V be the space of all continuous real valued functions $f : [0, 1] \longrightarrow \mathbb{R}$. Prove that the sets

$$B = \Big\{f \in V : |f(\tau)| \leq 1 \quad \text{for all} \quad \tau \in [0, 1]\Big\} \qquad \text{and}$$
$$K = \Big\{f \in V : f(\tau) \leq 0 \quad \text{for all} \quad \tau \in [0, 1]\Big\}$$

are convex.

(1.5) Some interesting properties. Some properties of convex sets are obvious, some are not so obvious, some are quite surprising. Here are some examples.

PROBLEMS.

1°. Prove that the intersection $\bigcap_{i \in I} A_i$ of convex sets is convex.

2°. Let $A \subset V$ be a convex set and let $T : V \longrightarrow W$ be a linear transformation. Prove that the image $T(A)$ is a convex set in W.

3. Let $A \subset \mathbb{R}^n$ be a polyhedron (see Problem 2, Section 1.3) and let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Prove that the image $T(A)$ is a polyhedron in $\mathbb{R}^m$.

4. Let $A, B \subset V$ be sets. Their *Minkowski sum* $A + B$ is a subset in V defined by

$$A + B = \Big\{x + y : x \in A, y \in B\Big\}.$$

For a number α and a subset $X \subset V$, let

$$\alpha X = \Big\{\alpha x : x \in X\Big\}.$$

Prove that $A + B$ is a convex set provided A and B are convex. Prove that for a convex set A and non-negative numbers α and β one has $(\alpha + \beta)A = \alpha A + \beta A$. Show that the identity does not hold if A is not convex or if α or β are allowed to be negative.

5*. For a set $A \subset \mathbb{R}^d$, let $[A] : \mathbb{R}^d \longrightarrow \mathbb{R}$ be the *indicator function* of A:

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

Let $A_1, \ldots, A_k$ be convex compact sets in $\mathbb{R}^n$ and let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Let $B_i = T(A_i)$ be the image of A_i. Suppose that $\sum_{i=1}^k \alpha_i [A_i] = 0$ for some numbers α_i. Prove that $\sum_{i=1}^k \alpha_i [B_i] = 0$. Show that this is no longer true if A_i are not convex.

6. Let $A \subset \mathbb{R}^d$ be a convex compact set. Let $B = \{(-1/d)x : x \in A\}$. Prove that there exists a vector $b \in \mathbb{R}^d$ such that $b+B \subset A$, where $b+B = \{b+x : x \in B\}$ is the translation of the set B by b.

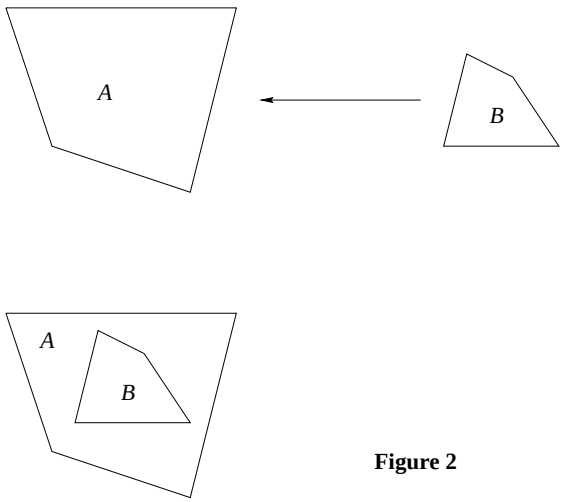


Figure 2

2. Properties of the Convex Hull. Carathéodory's Theorem

Recall (Definition 1.2) that the convex hull $\operatorname{conv}(S)$ of a set S is the set of all convex combinations of points from S.

(2.1) Theorem. *Let V be a vector space. The convex hull of $S \subset V$ is a convex set and any convex set containing S contains also* $\operatorname{conv}(S)$. *In other words,* $\operatorname{conv}(S)$ *is the smallest convex set containing S.*

Proof. First, we prove that $\operatorname{conv}(S)$ is a convex set itself (cf. Problem 1, Section 1.3). Indeed, let us choose two convex combinations $u = \alpha_1 u_1 + \ldots + \alpha_m u_m$ and $v = \beta_1 v_1 + \ldots + \beta_n v_n$ of points from S. The interval $[u, v]$ consists of the points $\gamma u + (1-\gamma)v$ for $0 \le \gamma \le 1$. Each such a point $\gamma\alpha_1 u_1 + \ldots + \gamma\alpha_m u_m + (1-\gamma)\beta_1 v_1 + \ldots + (1-\gamma)\beta_n v_n$ is a convex combination of points $u_1, \ldots, u_m, v_1, \ldots, v_n$ from S (since $\gamma\alpha_1 + \ldots + \gamma\alpha_m + (1-\gamma)\beta_1 + \ldots + (1-\gamma)\beta_n = \gamma(\alpha_1 + \ldots + \alpha_m) + (1-\gamma)(\beta_1 + \ldots + \beta_n) = \gamma + (1-\gamma) = 1$). Therefore, $\operatorname{conv}(S)$ is convex.

Now we have to prove that for any convex set A such that $S \subset A$, we must have $\operatorname{conv}(S) \subset A$. Let us choose a convex combination

$$u = \alpha_1 u_1 + \ldots + \alpha_m u_m$$

of points $u_1, \ldots, u_m$ from S. We must prove that $u \in A$. Without loss of generality, we may assume that $\alpha_i > 0$ for $i = 1, \ldots, m$. We proceed by induction on m. If $m = 1$ then $u = u_1$ and $u \in A$ since $S \subset A$. Suppose that $m > 1$. Then $\alpha_m < 1$ and we may write

$$u = (1-\alpha_m)w + \alpha_m u_m, \quad \text{where} \quad w = \frac{\alpha_1}{1-\alpha_m}u_1 + \ldots + \frac{\alpha_{m-1}}{1-\alpha_m}u_{m-1}.$$

Now, w is a convex combination of $u_1, \ldots, u_{m-1}$, because

$$\sum_{i=1}^{m-1} \frac{\alpha_i}{1-\alpha_m} = \frac{1}{1-\alpha_m} \sum_{i=1}^{m-1} \alpha_i = \frac{1-\alpha_m}{1-\alpha_m} = 1.$$

Therefore, by the induction conjecture, we have $w \in A$. Since A is convex, $[w, u_m] \subset A$, so $u \in A$. $\square$

PROBLEMS.

1°. Prove that $\mathrm{conv}\big(\mathrm{conv}(S)\big) = \mathrm{conv}(S)$ for any $S \subset V$.

2°. Prove that if $A \subset B$ then $\mathrm{conv}(A) \subset \mathrm{conv}(B)$.

3°. Prove that $\Big(\mathrm{conv}(A) \cup \mathrm{conv}(B)\Big) \subset \mathrm{conv}(A \cup B)$.

4. Let $S \subset V$ be a set and let $u, v \in V$ be points such that $u \notin \mathrm{conv}(S)$ and $v \notin \mathrm{conv}(S)$. Prove that if $u \in \mathrm{conv}\big(S \cup \{v\}\big)$ and $v \in \mathrm{conv}\big(S \cup \{u\}\big)$ then $u = v$.

5 (Gauss' Theorem). Let $f(z)$ be a polynomial in one complex variable z and let $z_1, \ldots, z_m$ be the roots of f (that is, the set of all solutions to the equation $f(z) = 0$). Let us interpret a complex number $z = x + iy$ as a point $(x, y) \in \mathbb{R}^2$. Prove that each root of the derivative $f'(z)$ lies in the convex hull $\mathrm{conv}(z_1, \ldots, z_m)$.

Hint: Without loss of generality we may suppose that $f(z) = (z - z_1) \cdots (z - z_m)$. If w is a root of $f'(z)$ then $\sum_{i=1}^{m} \prod_{j \neq i} (w - z_j) = 0$, and, therefore, $\sum_{i=1}^{m} \prod_{j \neq i} (\overline{w - z_j}) = 0$, where $\overline{z}$ is the complex conjugate of z. Multiply both sides of the last identity by $(w - z_1) \cdots (w - z_n)$ and express w as a convex combination of $z_1, \ldots, z_m$.

(2.2) Definition. The convex hull of a finite set of points in $\mathbb{R}^d$ is called a *polytope*.

PROBLEMS.

1. Prove that the set

$$\Delta = \Big\{ x \in \mathbb{R}^d : \xi_1 + \ldots + \xi_d = 1 \quad \text{and} \quad \xi_i \geq 0 : \ i = 1, \ldots, d \Big\}$$

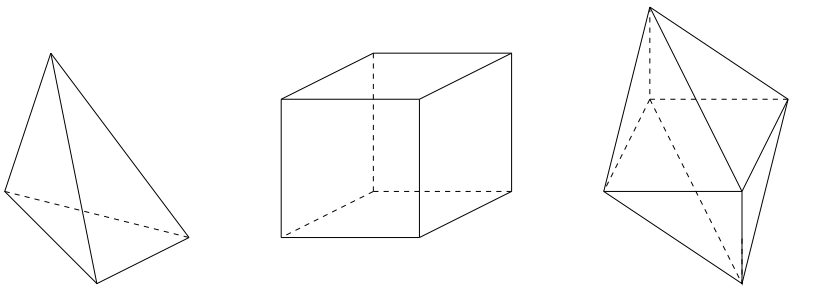
is a polytope in $\mathbb{R}^d$. This polytope is called a *simplex*.

2. Prove that the set

$$I = \Big\{ x \in \mathbb{R}^d : 0 \leq \xi_i \leq 1 : \ i = 1, \ldots, d \Big\}$$

is a polytope. This polytope is called a *cube*.

3. Prove that the set $O = \Big\{ x \in \mathbb{R}^d : |\xi_1| + \ldots + |\xi_d| \leq 1 \Big\}$ is a polytope. This polytope is called a *(hyper)octahedron* or *crosspolytope*.



simplex (tetrahedron) cube octahedron (crosspolytope)

some 3-dimensional polytopes

Figure 3

4. Prove that the disc $B = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1^2 + \xi_2^2 \leq 1\}$ is not a polytope.

5. Let $V = C[0,1]$ be the space of all real-valued continuous functions on the interval $[0,1]$ and let $A = \Big\{f \in V : 0 \leq f(\tau) \leq 1 \text{ for all } \tau \in [0,1]\Big\}$. Prove that A is not a polytope.

Recall from Problem 2, Section 1.3, that a polyhedron P is the set of solutions to a system of finitely many linear inequalities $P = \{x \in \mathbb{R}^d : \langle c_i, x\rangle \leq \alpha_i \text{ for } i = 1, \ldots, m\}$. The following two problems constitute the Weyl-Minkowski Theorem.

6*. Prove that a polytope $P \subset \mathbb{R}^d$ is also a polyhedron.

7*. Prove that a bounded polyhedron $P \subset \mathbb{R}^d$ is also a polytope.

It seems intuitively obvious that in the space of a small dimension we would need only few points to represent a given point from the convex hull as a convex combination. For example, in the plane $\mathbb{R}^2$ we would need only 3 points (see Figure 4). Thus to represent u as a convex combination, we need only 3 points (say, b, e and d).

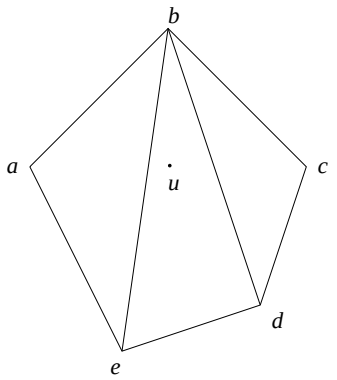


Figure 4

The general fact is known as Carathéodory's Theorem.

(2.3) Carathéodory's Theorem. *Let $S \subset \mathbb{R}^d$ be a set. Then every point $x \in \mathrm{conv}(S)$ can be represented as a convex combination of $d+1$ points from S:*

$$x = \alpha_1 y_1 + \ldots + \alpha_{d+1} y_{d+1}, \quad \textit{where} \quad \sum_{i=1}^{d+1} \alpha_i = 1, \quad \alpha_i \geq 0$$

and $y_i \in S$ for $i = 1, \ldots, d+1$.

Proof. Every point $x \in \mathrm{conv}(S)$ can be written as a convex combination

$$x = \alpha_1 y_1 + \ldots + \alpha_m y_m$$

of some points $y_1, \ldots, y_m \in S$. We can assume that $\alpha_i > 0$ for all $i = 1, \ldots, m$. If $m < d+1$, we can add terms $0y_1$, say, to get a convex combination with $d+1$ terms. Suppose that $m > d+1$. Let us show that we can construct a convex combination with fewer terms. Let us consider a system of linear homogeneous equations in m real variables $\gamma_1, \ldots, \gamma_m$:

$$\gamma_1 y_1 + \ldots + \gamma_m y_m = 0 \quad \text{and} \quad \gamma_1 + \ldots + \gamma_m = 0.$$

The first vector equation reads as d real linear equations

$$\gamma_1 \eta_{1j} + \ldots + \gamma_m \eta_{mj} = 0 : \quad j = 1, \ldots, d,$$

where η_{ij} are the coordinates of y_i: $y_i = (\eta_{i1}, \ldots, \eta_{id})$. Altogether, we have $d+1$ linear homogeneous equations in m variables $\gamma_1, \ldots, \gamma_m$. Since $m > d+1$, there must be a non-trivial solution $\gamma_1, \ldots, \gamma_m$. Since $\gamma_1 + \ldots + \gamma_m = 0$, some γ_i are strictly positive and some are strictly negative. Let

$$\tau = \min\{\alpha_i/\gamma_i : \gamma_i > 0\} = \alpha_{i_0}/\gamma_{i_0}.$$

Let $\widetilde{\alpha}_i = \alpha_i - \tau\gamma_i$ for $i = 1, \ldots, m$. Then $\widetilde{\alpha}_i \geq 0$ for all $i = 1, \ldots, m$ and $\alpha_{i_0} = 0$. Furthermore,

$$\widetilde{\alpha}_1 + \ldots + \widetilde{\alpha}_m = (\alpha_1 + \ldots + \alpha_m) - \tau(\gamma_1 + \ldots + \gamma_m) = 1$$

and

$$\widetilde{\alpha}_1 y_1 + \ldots + \widetilde{\alpha}_m y_m = \alpha_1 y_1 + \ldots + \alpha_m y_m - \tau(\gamma_1 y_1 + \ldots + \gamma_m y_m) = x$$

Therefore, we represented x as a convex combination

$$x = \sum_{i \neq i_0} \widetilde{\alpha}_i y_i$$

of $m - 1$ points $y_1, \ldots, \widehat{y_{i_0}}, \ldots, y_m$ (y_{i_0} omitted).

So, if x is a convex combination of $m > d + 1$ points, it can be written as a convex combination of fewer points. Iterating this procedure, we get x as a convex combination of $d + 1$ (or fewer) points from S. □

PROBLEMS.

1°. Show by an example that the constant $d + 1$ in Carathéodory's theorem can not be improved to d.

2*. Let $S_1, \ldots, S_{d+1}$ be subsets of $\mathbb{R}^d$. Prove that if $u \in \mathrm{conv}(S_i)$ for each S_i, then there exist points $v_i \in S_i$ such that $u \in \mathrm{conv}\big(v_1, \ldots, v_{d+1}\big)$.

Hint: Choose points $v_i \in S_i$ in such a way that the distance from u to $\mathrm{conv}(v_1, \ldots, v_{d+1})$ is the minimal possible. Prove that if $u \notin \mathrm{conv}\big(v_1, \ldots, v_{d+1}\big)$, the distance could have been decreased further.

3. Let $S \subset \mathbb{R}^d$ be a set and let u be a point *in the interior* of $\mathrm{conv}(S)$. Then one can choose $2d$ points $v_1, \ldots, v_{2d} \in S$ such that u lies *in the interior* of $\mathrm{conv}\big(v_1, \ldots, v_{2d}\big)$.

4*. Suppose that $S \subset \mathbb{R}^d$ is a connected set (or even a union of d connected sets). Prove that every point $u \in \mathrm{conv}(S)$ is a convex combination of some d points of S.

(2.4) Corollary. *Suppose that $S \subset \mathbb{R}^d$ is a compact set. Then $\mathrm{conv}(S)$ is a compact set.*

Proof. Let $\Delta \subset \mathbb{R}^{d+1}$ be the set

$$\Delta = \Big\{(\alpha_1, \ldots, \alpha_{d+1}) : \sum_{i=1}^{d+1} \alpha_i = 1 \quad \text{and} \quad \alpha_i \geq 0 \text{ for } i = 1, \ldots, d+1\Big\}.$$

Then Δ is compact and so is

$$S^{d+1} \times \Delta = \Big\{\big(u_1, \ldots, u_{d+1}; \alpha_1, \ldots, \alpha_{d+1}\big) : u_i \in S; (\alpha_1, \ldots, \alpha_{d+1}) \in \Delta\Big\}.$$

Consider the map $\Phi : S^{d+1} \times \Delta \longrightarrow \mathbb{R}^d$, $\Phi(u_1, \ldots, u_{d+1}; \alpha_1, \ldots, \alpha_{d+1}) = \alpha_1 u_1 + \ldots + \alpha_{d+1} u_{d+1}$. Theorem 2.3 implies that the image of Φ is $\mathrm{conv}(S)$. Since Φ is continuous, the proof follows. $\square$

PROBLEMS.

1. Give an example of a closed set in $\mathbb{R}^2$ whose convex hull is not closed.
2. Prove that the convex hull of an open set in $\mathbb{R}^d$ is open.

3. Theorems of Radon and Helly

The following very useful and, essentially, simple result was first stated 1921 by J. Radon as a lemma.

(3.1) Radon's Theorem. *Let $S \subset \mathbb{R}^d$ be a set containing at least $d+2$ points. Then there are two non-intersecting subsets $R, B \subset S$: $R \cap B = \emptyset$ such that*

$$\mathrm{conv}(R) \cap \mathrm{conv}(B) \neq \emptyset.$$

Proof. Let $v_1, \ldots, v_m$, $m \geq d+2$ be distinct points from S. Consider the following system of $d+1$ homogeneous linear equations in variables $\gamma_1, \ldots, \gamma_m$:

$$\gamma_1 v_1 + \ldots + \gamma_m v_m = 0, \quad \gamma_1 + \ldots + \gamma_m = 0.$$

Since $m \geq d+2$, there is a non-trivial solution to this system. Let

$$R = \big\{v_i : \ \gamma_i > 0\big\} \quad \text{and} \quad B = \big\{v_i : \ \gamma_i < 0\big\}.$$

Then $R \cap B = \emptyset$.

Let $\beta = \sum_{i:\gamma_i>0} \gamma_i$. Then $\beta > 0$ and $\sum_{i:\gamma_i<0} \gamma_i = -\beta$, since γ's sum up to zero. Since $\gamma_1 v_1 + \ldots + \gamma_m v_m = 0$, we have

$$\sum_{i:\gamma_i>0} \gamma_i v_i = \sum_{i:\gamma_i<0} (-\gamma_i) v_i.$$

Let

$$v = \sum_{i:\gamma_i>0} \frac{\gamma_i}{\beta} v_i = \sum_{i:\gamma_i<0} \frac{-\gamma_i}{\beta} v_i.$$

It is seen that v is a convex combination of the points from R and a convex combination of the points from B. In other words, $v \in \mathrm{conv}(R)$ and $v \in \mathrm{conv}(B)$. $\square$

For example, for any set of four points in the plane, either one of them lies within the convex hull of the other three, or they can be split into two pairs whose convex hulls intersect.

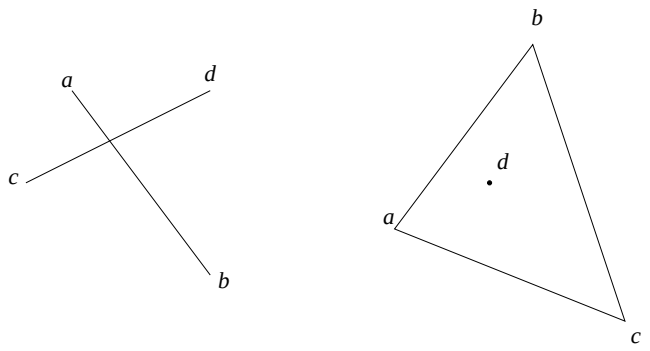


Figure 5

PROBLEMS.

1°. Show by an example that the constant $d+2$ in Radon's Theorem can not be improved to $d+1$.

2* (Tverberg's Theorem). Let $k \geq 2$. Prove that for any set S of $(k-1)(d+1)+1$ or more points in $\mathbb{R}^d$, one can find k pairwise non-intersecting subsets $A_1, \ldots, A_k \subset S$ such that the intersection

$$\mathrm{conv}(A_1) \cap \mathrm{conv}(A_2) \cap \ldots \cap \mathrm{conv}(A_k)$$

is not empty. Show that for some sets of $(k-1)(d+1)$ points, such subsets $A_1, \ldots, A_k$ can not be found.

The following result (probably, the most famous result in convexity) was discovered by E. Helly in 1913. The proof below belongs to Radon (1921).

(3.2) Helly's Theorem. *Let $A_1, \ldots, A_m$ be a finite family of convex sets in $\mathbb{R}^d$. Suppose that every $d+1$ of the sets have a common point:*

$$A_{i_1} \cap \ldots \cap A_{i_{d+1}} \neq \emptyset.$$

Then all the sets have a common point:

$$A_1 \cap \ldots \cap A_m \neq \emptyset.$$

Proof. The proof is by induction on m (starting with $m = d+1$). Suppose that $m > d+1$. Then, by the induction conjecture, for every $i = 1, \ldots, m$ there is a point p_i in the intersection $A_1 \cap \ldots \cap A_{i-1} \cap A_{i+1} \cap \ldots \cap A_m$ (A_i is missing). Altogether, we have $m > d+1$ points p_i, each of which belongs to all the sets, except, perhaps, A_i. If two of these points happened to coincide, we get a point which belongs to all A_i-s. Otherwise, by Radon's Theorem (Theorem 3.1) there are non-intersecting subsets $R = \{p_i : i \in I\}$ and $B = \{p_j : j \in J\}$ such that there is a point

$$p \in \mathrm{conv}(R) \cap \mathrm{conv}(B).$$

We claim that p is the common point of $A_1, \ldots, A_m$. Indeed, all the points $p_i : i \in I$ of R belong to the sets $A_i : i \notin I$. All the points $p_j : j \in J$ of B belong to the sets $A_j : j \notin J$. Since the sets A_i are convex, every point from $\mathrm{conv}(R)$ belongs to the sets $A_i : i \notin I$. Similarly, every point from $\mathrm{conv}(B)$ belongs to the sets $A_j : j \notin J$. Therefore,

$$p \in \bigcap_{i \notin I} A_i \quad \text{and} \quad p \in \bigcap_{j \notin J} A_j.$$

Since $I \cap J = \emptyset$ we have

$$p \in \bigcap_{i=1}^{m} A_i$$

and the proof follows. □

PROBLEMS.

1°. Show that the theorem does not hold for non-convex sets A_i.

2°. Construct an example of convex sets A_i in $\mathbb{R}^2$, such that every two sets have a common point, but there is no point which would belong to all sets A_i.

3°. Give an example of an infinite family $\{A_i : i = 1, 2, \ldots\}$ of convex sets in $\mathbb{R}^d$ such that every $d+1$ sets have a common point but there isn't a point common to all sets A_i.

The theorem can be extended to infinite families of convex compact sets

(3.3) Corollary. *Let $\{A_i : i \in I\}$ be a (possibly infinite) family of convex compact sets in $\mathbb{R}^d$ such that the intersection of any $d+1$ sets is not empty:*

$$A_{i_1} \cap \ldots \cap A_{i_{d+1}} \neq \emptyset.$$

Then the intersection of all the sets A_i is not empty:

$$\bigcap_{i \in I} A_i \neq \emptyset.$$

Proof. By Theorem 3.2, for any finite subfamily $J \subset I$ the intersection $\bigcap_{i \in J} A_i$ is not empty. We use the fact that if the intersection of a family of compact sets is empty, then the intersection of the sets from some finite subfamily is empty. □

Helly's Theorem has numerous generalizations, extensions, ramifications, *etc.* To list all of them is impossible, here are just some.

PROBLEMS.

1. Let $A_1, \ldots, A_m$ be convex sets in $\mathbb{R}^d$ and let $k \leq d+1$. Prove that, if every k of the sets have a point in common, then for every $(d-k+1)$-dimensional subspace L in $\mathbb{R}^d$ there exists a translate $L + u : u \in \mathbb{R}^d$ which intersects every set $A_i : i = 1, \ldots, m$.

2. Let $A_1, \ldots, A_m$ and C be convex sets in $\mathbb{R}^d$. Suppose that for any $d+1$ sets $A_{i_1}, \ldots, A_{i_{d+1}}$ there is a translate $C+u : u \in \mathbb{R}^d$ of C which *intersects* all $A_{i_1}, \ldots, A_{i_{d+1}}$. Prove that there is a translate $C+u$ of C which *intersects* all sets $A_1, \ldots, A_m$.

3. In Problem 2, replace *intersects* by *contains*.

4. In Problem 2, replace *intersects* by *is contained in*.

5* (Fractional Helly's Theorem). Let $A_1, \ldots, A_m$ be convex sets in $\mathbb{R}^d$. Let f be the number of $(d+1)$-subfamilies $A_{i_1}, \ldots, A_{i_{d+1}}$ that have a point in common. Let us fix $0 \leq \rho \leq 1$. Prove that if $f > \left(1-(1-\rho)^{d+1}\right)\binom{m}{d+1}$, then some $\lfloor \rho m \rfloor + 1$ sets A_i have a point in common.

6* ("Piercing" Theorem). Prove that for every triple (p, q, d), such that $p \geq q \geq d+1$ there exists a positive integer $c(p, q, d)$ such that if $A_1, \ldots, A_m \subset \mathbb{R}^d$ are convex sets, $m \geq p$ and out of every p sets $A_{i_1}, \ldots, A_{i_p}$ some q have a point in common, then some set X of $c(p, q, d)$ points in $\mathbb{R}^d$ intersects every set A_i.

7* (Measure of the intersection). Let $A_1, \ldots, A_m \subset \mathbb{R}^d$ be convex sets. Suppose that $m \geq 2d$ and that the intersection of every $2d$ sets $A_{i_1}, \ldots, A_{i_{2d}}$ has volume at least 1. Then the intersection $A_1 \cap \ldots \cap A_m$ has volume at least d^{-2d^2} (it is conjectured that the constant d^{-2d^2} can be improved to $d^{-\alpha d}$ for some absolute constant $\alpha > 0$).

4. Some Applications of Helly's Theorem

(4.1) Separating points by a hyperplane. Suppose that there is a finite set R of red points in $\mathbb{R}^d$ and a finite set B of blue points in $\mathbb{R}^d$. A *hyperplane* $H \subset \mathbb{R}^d$ is the set described by a linear equation $H = \left\{x \in \mathbb{R}^d : \langle c, x\rangle = \alpha\right\}$, where $c \neq 0$ is a non-zero vector and α is a number. We say that the hyperplane $H \subset \mathbb{R}^d$ strictly separates red and blue points if $\langle c, x\rangle < \alpha$ for all $x \in R$ and $\langle c, x\rangle > \alpha$ for all $x \in B$. The following result was proved by P. Kirchberger in 1903, that is, *before* Helly's Theorem.

Proposition. *Suppose that for any set $S \subset \mathbb{R}^d$ of $d+2$ points (or fewer), there exists a hyperplane which strictly separates the sets $S \cap R$ and $S \cap B$ of red, resp. blue points in S. Then there exists a hyperplane which strictly separates the sets R and B.*

Proof. A hyperplane $H = \left\{x : \langle c, x\rangle = \alpha\right\}$, where $c = (\gamma_1, \ldots, \gamma_d) \in \mathbb{R}^d$ can be encoded by a point $(c, \alpha) = (\gamma_1, \ldots, \gamma_d, \alpha) \in \mathbb{R}^{d+1}$. For every point $r \in R$ we define a set $A_r \subset \mathbb{R}^{d+1}$:

$$A_r = \left\{(c, \alpha) \in \mathbb{R}^{d+1} : \langle c, r\rangle < \alpha\right\}$$

and for every point $b \in B$ we define a set $A_b \subset \mathbb{R}^{d+1}$:

$$A_b = \left\{(c, \alpha) \in \mathbb{R}^{d+1} : \langle c, b\rangle > \alpha\right\}.$$

It is clear that A_b and A_r are convex sets in $\mathbb{R}^{d+1}$. Therefore, by Helly's Theorem the intersection

$$\Big(\bigcap_{r\in R} A_r\Big) \cap \Big(\bigcap_{b\in B} A_b\Big)$$

is non-empty, provided for any subset $S \subset R \cup B$ of $d+2$ points the intersection

$$\Big(\bigcap_{r\in S\cap R} A_r\Big) \cap \Big(\bigcap_{b\in S\cap B} A_b\Big)$$

is non-empty. Since the sets A_r and A_b are open, their intersection is an open set; therefore an intersection of sets A_r and A_b is non-empty if and only if it contains a point (c, α) with $c \neq 0$. Hence for any subset $S \subset R \cup B$, the intersection

$$\Big(\bigcap_{r\in S\cap R} A_r\Big) \cap \Big(\bigcap_{b\in S\cap B} A_b\Big)$$

is not empty if and only if there is a point $(c, \alpha) \in \mathbb{R}^{d+1}$ such that the hyperplane

$$H = \big\{x \in \mathbb{R}^d : \langle c, x\rangle = \alpha\big\}$$

strictly separates the sets $B \cap S$ and $R \cap S$ The proof now follows. □

For example, in the plane $\mathbb{R}^2$, a set of red points can not be separated from a set of blue points if and only if red and blue points in some four-point subset can not be separated.

The two reasons points can not be separated in the plane

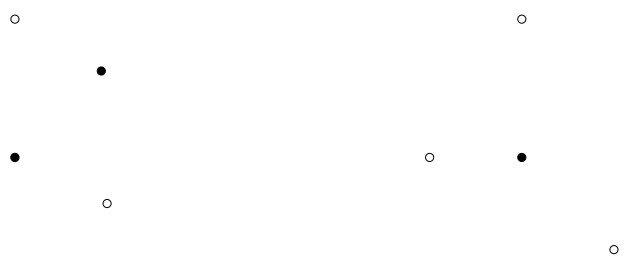


Figure 6

PROBLEMS.

1. Prove that if a convex set is contained in the union of a finite family of halfspaces in $\mathbb{R}^d$ (sometimes we say *covered* by a finite family of halfspaces), then

it is contained in the union of some $d+1$ (or fewer) halfspaces from the family (covered by some $d+1$ subspaces).

2. Let $I_1, \ldots, I_m$ be parallel line segments in $\mathbb{R}^2$, such that for every three $I_{i_1}, I_{i_2}, I_{i_3}$ there is a straight line that intersects all the three. Then there is a straight line that intersects all the segments $I_1, \ldots, I_m$.

3. Let $A_i : i = 1, \ldots, m$ be convex sets in $\mathbb{R}^2$ such that for every two sets A_i and A_j there is a line parallel to x-axis which intersects them both. Prove that there is a line parallel to x-axis which intersects all sets A_i.

(4.2) The center point. Let us fix a Borel probability measure μ on $\mathbb{R}^d$. The following particular cases are of interest:

Counting measure. Suppose that there is a finite set $X \subset \mathbb{R}^d$ of $|X| = n$ points and $\mu(A) = |A \cap X|/n$ is the proportion of the points in X contained in A.

Integrable density. Suppose that there is an integrable function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that $f(x) \geq 0$ for all $x \in \mathbb{R}^d$ and such that $\int_{\mathbb{R}^d} f(x)\, dx = 1$, where dx is the Lebesgue measure. Let $\mu(A) = \int_A f(x)\, dx$ for all (Borel) measurable sets A.

With a hyperplane $H = \{x : \langle c, x\rangle = \alpha\}$ we associate two *open halfspaces*

$$H_+ = \{x : \langle c, x\rangle > \alpha\} \quad \text{and} \quad H_- = \{x : \langle c, x\rangle > \alpha\}$$

and two closed subspaces

$$\overline{H_+} = \{x : \langle c, x\rangle \geq \alpha\} \quad \text{and} \quad \overline{H_-} = \{x : \langle c, x\rangle \leq \alpha\}.$$

Proposition. *Let μ be a Borel probability measure on $\mathbb{R}^d$. Then there exists a point $y \in \mathbb{R}^d$ such that for any closed halfspace $\overline{H_\star}$ containing y one has*

$$\mu(\overline{H_\star}) \geq \frac{1}{d+1}.$$

Proof. For a closed halfspace $G \subset \mathbb{R}^d$, let $\widetilde{G} = \mathbb{R}^d \setminus G$ be the complementary open halfspace. Let $\mathcal{S}$ be the set of all closed halfspaces G such that $\mu\big(\widetilde{G}\big) < 1/(d+1)$. We observe that for any $d+1$ halfspaces $G_1, \ldots, G_{d+1}$ from $\mathcal{S}$ one has

$$\mu\big(\widetilde{G}_1 \cup \ldots \cup \widetilde{G}_{d+1}\big) < \frac{(d+1)}{(d+1)} = 1 \quad \text{and hence} \quad \widetilde{G}_1 \cup \ldots \cup \widetilde{G}_{d+1} \neq \mathbb{R}^d,$$

which implies that $G_1 \cap \ldots \cap G_{d+1} \neq \emptyset$. Helly's Theorem implies that any finite family $\{G_i\}$ of halfspaces from $\mathcal{S}$ has a non-empty intersection. We can choose a finite number of halfspaces $G_1, \ldots, G_m$ from $\mathcal{S}$ such that the intersection $B = G_1 \cap \ldots \cap G_m$ is bounded and hence compact. Thus $\{B \cap G : G \in \mathcal{S}\}$ is a family of

compact sets such that every finite subfamily has a non-empty intersection. Hence there is a point y which belongs to all halfspaces $G \in \mathcal{S}$. It follows then that for any open halfspace $H_\star$ containing y we must have $\mu(H_\star) \geq 1/(d+1)$. Since μ is σ-additive and a closed halfspace can be represented as an intersection of countably many nested open halfspaces, the proof follows. $\square$

The result was obtained in 1916 by J. Radon in a stronger form (one can choose y to be the center of mass). The above proof belongs to I.M. Yaglom and V.G. Boltyanskii (1956).

PROBLEMS.

1. Let $S \subset \mathbb{R}^d$ be a convex compact set. Prove that there is a point $u \in \mathbb{R}^d$ such that $(-1/d)S + u \subset S$. Here, as usual, $\alpha S = \{\alpha x : x \in S\}$ and $S + y = \{x + y : x \in S\}$.

Hint: For every point $x \in S$ consider the set $A_x = \{u : (-1/d)x + u \in S\}$. Use Helly's Theorem.

2* ("Ham Sandwich Theorem"). Let $\mu_1, \ldots, \mu_d$ be a set of Borel probability measures on $\mathbb{R}^d$. Prove that there exists a hyperplane $H \subset \mathbb{R}^d$ such that $\mu_i(\overline{H_+}) \geq 1/2$ and $\mu_i(\overline{H_-}) \geq 1/2$ for all $i = 1, \ldots, d$.

3* (Center Transversal Theorem). Let $\mu_1, \ldots, \mu_k$, $k \leq d$ be Borel probability measures on $\mathbb{R}^d$. Prove that there exists a $(k-1)$-dimensional affine subspace $L \subset \mathbb{R}^d$ such that for every closed halfspace $\overline{H_\star}$ containing L we have $\mu_i\big(\overline{H_\star}\big) \geq 1/(d-k+2)$.

4°. Prove that the theorem of Problem 3 implies both the theorem of problem 2 and Proposition of this section.

A couple of geometric problems.

5* (Krasnoselsky's Theorem). Let $X \subset \mathbb{R}^d$ be a set and let $a, b \in X$ be points. We say that b is *visible* from a if $[a, b] \subset X$. Suppose that $X \subset \mathbb{R}^d$ is an infinite compact set such that for any $(d+1)$ points of X there is a point from which all $d+1$ are visible. Prove that there is a point from which all points of X are visible.

6 (Jung's Theorem). For a compact set $X \subset \mathbb{R}^d$, let us call $\max_{y,z \in X} \|y - z\|$ *diameter* of X. Prove that a compact set of diameter 2 is contained in a ball of radius $\sqrt{2d/(d+1)}$.

5. An Application to Approximation

(5.1) Uniform approximations.

Let us fix some real-valued functions $\phi_i : X \longrightarrow \mathbb{R}$, $i = 0, \ldots, d$ on some set X. Given a function $f : X \longrightarrow \mathbb{R}$ and a number $\epsilon \geq 0$, we want to construct a linear combination $\Phi_a(x) = \alpha_0\phi_0(x) + \alpha_1\phi_1(x) + \ldots + \alpha_d\phi_d(x)$: $a = (\alpha_0, \ldots, \alpha_d)$ such that

$$|f(x) - \Phi_a(x)| \leq \epsilon \quad \text{for all } x \in X.$$

This is the problem of the *uniform* or *Chebyshev* approximation.

(5.2) Proposition. *Suppose that X is a finite set. Let us fix an $\epsilon \geq 0$. Suppose that for any $d+2$ points $x_{i_1}, \ldots, x_{i_{d+2}}$ from X one can construct a function $\Phi_a(x)$ (with $a = (\alpha_0, \ldots, \alpha_d)$ depending on $x_{i_1}, \ldots, x_{i_{d+2}}$) such that*

$$|f(x) - \Phi_a(x)| \leq \epsilon \quad \text{for } x = x_{i_1}, \ldots, x_{i_{d+2}}.$$

Then there exists a function $\Phi_a(x)$ such that

$$|f(x) - \Phi_a(x)| \leq \epsilon \quad \text{for all } \ x \in X.$$

Proof. For an $x \in X$ let us define a set $A(x) \subset \mathbb{R}^{d+1}$:

$$A(x) = \Big\{(\alpha_0, \ldots, \alpha_d) : \big|f(x) - \alpha_0\phi_0(x) - \ldots - \alpha_d\phi_d(x)\big| \leq \epsilon\Big\}.$$

In other words, $A(x)$ is the set of functions Φ_a that approximate f within ϵ at the point x. We observe that $A(x)$ are convex sets (see Problem 1° below) and that

$$A(x_{i_1}) \cap \ldots \cap A(x_{i_{d+2}}) \neq \emptyset$$

for all possible choices of $d+2$ points $x_{i_1}, \ldots, x_{i_{d+2}}$ in X. Since X is finite, Helly's Theorem (Theorem 3.2) implies that the intersection of all sets $A(x) : x \in X$ is non-empty. It follows then that for a point

$$a \in \bigcap_{x \in X} A(x)$$

we have $|f(x) - \Phi_a(x)| \leq \epsilon$ for all $x \in X$. □

PROBLEMS.

1°. Prove that $A(x) \subset \mathbb{R}^{d+1}$ is a closed convex set.

2°. Show that for $d \geq 1$ the set $A(x)$ is *not* compact.

To prove an analogue of Proposition 5.2 for infinite sets X we must assume some regularity of functions $\phi_0, \ldots, \phi_d$.

(5.3) Proposition. *Suppose there is a finite set of points $y_1, \ldots, y_m$ in X such that whenever for some $\Phi_a = \alpha_0\phi_0 + \ldots + \alpha_d\phi_d$ we have $\Phi_a(y_1) = \ldots = \Phi_a(y_m) = 0$, we must have $\alpha_0 = \ldots = \alpha_d = 0$. Suppose further, that for any set of $d+2$ points $x_1, \ldots, x_{d+2}$ of X, one can construct a function $\Phi_a(x)$ (with $a = (\alpha_0, \ldots, \alpha_d)$ depending on $x_1, \ldots, x_{d+2}$) such that*

$$|f(x) - \Phi_a(x)| \leq \epsilon \quad \text{for } \ x = x_1, \ldots, x_{d+2}.$$

Then there exists a function $\Phi_a(x)$ such that

$$|f(x) - \Phi_a(x)| \leq \epsilon \quad \text{for all } \ x \in X.$$

Proof. Let $A(x) : x \in X$ be the sets defined in the proof of Proposition 5.2. Let

$$A = A(y_1) \cap \ldots \cap A(y_m).$$

First, we prove that A is compact. Indeed, by Problem 1°, Section 5.2, the set A is closed. It remains to show that A is bounded. Let us define a function

$$G : \mathbb{R}^{d+1} \longrightarrow \mathbb{R}, \quad G(a) = \max\big\{|\Phi_a(x)| : x = y_1, \ldots, y_M\big\}.$$

Then $G(\lambda a) = |\lambda| G(a)$ for $\lambda \in \mathbb{R}$, $G(a) > 0$ for $a \neq 0$ and G is continuous (in fact, G is a norm on $\mathbb{R}^d$). Therefore,

$$\min\big\{G(a) : \|a\| = 1\big\} = \delta > 0$$

and $G(a) > \delta R$ provided $\|a\| > R$. Now, if $|f(x) - \Phi_a(x)| \leq \epsilon$ for $x = y_1, \ldots, y_M$, we have $|\Phi_a(x)| \leq |f(x)| + \epsilon$ for $x = y_1, \ldots, y_M$. Letting

$$R = \epsilon + \max\big\{|f(y_i)| : i = 1, \ldots, M\big\},$$

we conclude that $G(a) \leq R$, and, therefore, $\|a\| \leq R/\delta$ for any $a \in A$.

For $x \in X$ let $\overline{A(x)} = A(x) \cap A$. Then each set $\overline{A(x)}$ is compact. Applying Helly's Theorem as in the proof of Proposition 5.2, we conclude that every intersection of *finite* family of sets $A(x)$ is non-empty. Therefore, every intersection $\overline{A(x_1)} \cap \ldots \overline{A(x_{d+2})}$ is a non-empty compact convex set. Therefore, By Corollary 3.3, the intersection of all the sets $\overline{A(x)}$ is non-empty and so is the intersection of all the sets $A(x)$. A point

$$a = (\alpha_0, \ldots, \alpha_d) \in \bigcap_{x \in X} A(x)$$

gives rise to a function

$$\Phi_a = \alpha_0 \phi_0 + \ldots + \alpha_d \phi_d,$$

which approximates $f(x)$ uniformly within the error ϵ. □

PROBLEMS.

In the problems below, $X = [0, 1]$ and $\phi_i(x) = x^i$, $i = 0, \ldots, d$.

1°. Prove that for any $d + 1$ distinct points $y_1, y_2, \ldots, y_{d+1}$ from $[0, 1]$ the intersection $A(y_1) \cap \ldots \cap A(y_{d+1})$ is compact.

2°. Let $f(x) = e^x$ for $x \in [0, 1]$. Let us choose $\epsilon = 0$. Check that each intersection $A(x_1) \cap \ldots A(x_{d+1})$ is not empty for any choice of $x_1, \ldots, x_{d+1} \in [0, 1]$, but $\bigcap_{x \in [0,1]} A_x = \emptyset$. In other words, for every $d + 1$ points $x_1, \ldots, x_{d+1}$ there is a polynomial $p(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_d x^d$ such that $p(x) = e^x$ for $x = x_1, \ldots, x_{d+1}$ but there isn't a polynomial $p(x)$ such that $e^x = p(x)$ for all $x \in [0, 1]$.

3*. Let $f : [0, 1] \longrightarrow \mathbb{R}$ be any function. Prove that for any $d+2$ points $0 \leq x_1 < x_2 < \ldots < x_{d+2} \leq 1$ there is a unique polynomial $p(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_d x^d$, such that

$$|f(x_1) - p(x_1)| = |f(x_2) - p(x_2)| = \ldots = |f(x_{d+2}) - p(x_{d+2})|$$

and the signs of the differences

$$f(x_1) - p(x_1), f(x_2) - p(x_2), \ldots, f(x_{d+2}) - p(x_{d+2})$$

alternate. The polynomial p gives the unique best (that is, with the smallest ϵ) uniform approximation to f on the set of $d+2$ points $x_1, \ldots, x_{d+2}$. The error ϵ of this approximation can be found as $\epsilon = |\beta|$, where $\alpha_0, \ldots, \alpha_d, \beta$ is the (necessarily unique) solution to the system of $d+2$ linear equations

$$f(x_1) - p(x_1) = \beta,\ f(x_2) - p(x_2) = -\beta, \ldots, f(x_{d+2}) - p(x_{d+2}) = (-1)^{d+1}\beta$$

in $d+2$ variables $(\alpha_0, \ldots, \alpha_d, \beta)$.

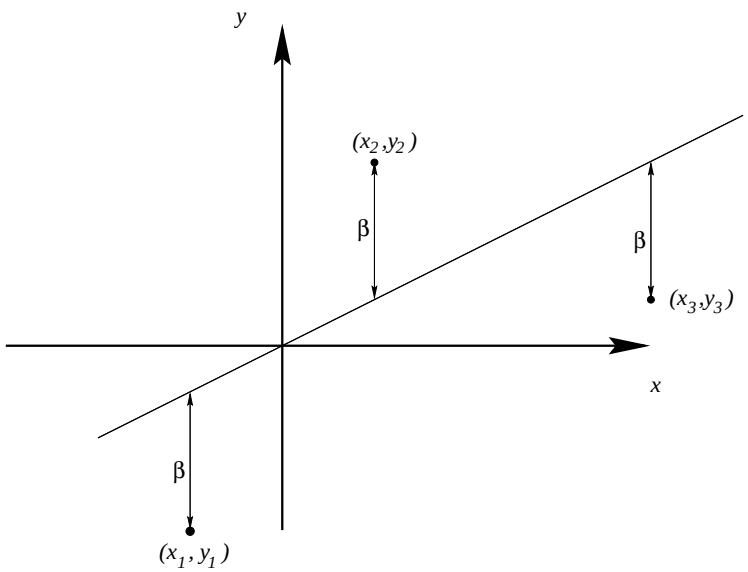


Figure 7

Figure 7 shows a linear function $p(x) = \alpha_0 + \alpha_1 x$ which provides the best uniform approximation for f at some three points x_1, x_2 and x_2 and satisfies $p(x_1) - f(x_1) = -\big(p(x_2) - f(x_2)\big) = p(x_3) - f(x_3)$.

6. Euler Characteristic

Helly's Theorem tells us something special about how convex sets may intersect. Here we introduce another powerful tool to study intersection properties of convex sets.

(6.1) Definition. Let $A \subset \mathbb{R}^d$ be a subset. The *indicator function* χ_A of A is the function $\chi_A : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that

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

Often, especially when we do not need to mention the argument x, we write $[A]$ instead of χ_A.

PROBLEM.

1°. Prove that $[A] \cdot [B] = [A \cap B]$.

(6.2) Lemma (Inclusion-Exclusion Formula). *Let $A_1, \ldots, A_m \subset \mathbb{R}^d$ be sets. Then*

$$[A_1 \cup \ldots \cup A_m] = 1 - (1 - [A_1]) \cdot (1 - [A_2]) \cdots (1 - [A_m])$$

$$= \sum_{k=1}^{m} (-1)^{k-1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le m} [A_{i_1} \cap \ldots \cap A_{i_k}].$$

In particular,

$$[A_1 \cup A_2] = [A_1] + [A_2] - [A_1 \cap A_2].$$

Proof. Let us choose an $x \in \mathbb{R}^d$. Then

$$\chi_{A_1 \cup \ldots \cup A_m}(x) = \begin{cases} 1 & \text{if } x \in A_1 \cup \ldots \cup A_m \\ 0 & \text{if } x \notin A_1 \cup \ldots \cup A_m. \end{cases}$$

On the other hand,

$$1 - \chi_{A_i}(x) = \begin{cases} 1 & \text{if } x \notin A_i \\ 0 & \text{if } x \in A_i. \end{cases}$$

Therefore,

$$\big(1 - \chi_{A_1}(x)\big) \cdots \big(1 - \chi_{A_m}(x)\big) = \begin{cases} 1 & \text{if } x \notin A_i \text{ for all } i \\ 0 & \text{if } x \in A_i \text{ for some } i. \end{cases}$$

Therefore, $[A_1 \cup \ldots \cup A_m] = 1 - (1 - [A_1]) \cdot (1 - [A_2]) \cdots (1 - [A_m])$. Expanding the product, we get the second formula. $\square$

PROBLEM.

1°. Researchers at some research institute speak French, Russian, and English. Among them 20 people speak French, 15 speak Russian, and 10 speak English. Also, 8 people speak French and Russian, 5 people speak Russian and English and 7 speak French and English. Two people speak French, Russian and English. How many people work at the institute?

(6.3) Definitions. The real vector space spanned by the functions $[A]$, where $A \subset \mathbb{R}^d$ is a convex compact set, is called the *algebra of convex compact sets* and denoted $\mathcal{K}(\mathbb{R}^d)$. Thus a function $f \in \mathcal{K}(\mathbb{R}^d)$ is a linear combination

$$f = \sum_{i=1}^{m} \alpha_i [A_i],$$

where $[A_i] \subset \mathbb{R}^d$ are convex compact sets and $\alpha_i \in \mathbb{R}$ are real numbers.

The real vector space spanned by the functions $[A]$, where $A \subset \mathbb{R}^d$ is a closed convex set is called the *algebra of convex closed sets* and denoted $\mathcal{C}(\mathbb{R}^d)$. Thus a typical function $f \in \mathcal{C}(\mathbb{R}^d)$ is a linear combination

$$f = \sum_{i=1}^{m} \alpha_i [A_i],$$

where $[A_i] \subset \mathbb{R}^d$ are convex closed sets and $\alpha_i \in \mathbb{R}$ are real numbers.

A linear functional $\nu : \mathcal{K}(\mathbb{R}^d) \longrightarrow \mathbb{R}$, resp. $\nu : \mathcal{C}(\mathbb{R}^d) \longrightarrow \mathbb{R}$, is called a *valuation*. Thus $\nu(\alpha f + \beta g) = \alpha\nu(f) + \beta\nu(g)$ for any real α and β and any $f, g \in \mathcal{K}(\mathbb{R}^d)$, resp. $f, g \in \mathcal{C}(\mathbb{R}^d)$. More generally, we call *valuation* any linear transformation $\mathcal{K}(\mathbb{R}^d), \mathcal{C}(\mathbb{R}^d) \longrightarrow V$, where V is a real vector space.

PROBLEMS.

1°. Prove that the product fg of functions $f, g \in \mathcal{K}(\mathbb{R}^d)$ is a function in $\mathcal{K}(\mathbb{R}^d)$ and that the product fg of functions $f, g \in \mathcal{C}(\mathbb{R}^d)$ is a function in $\mathcal{C}(\mathbb{R}^d)$.

2°. Do the functions $[A] : A \subset \mathbb{R}^d$ is a non-empty convex compact set form a basis of $\mathcal{K}(\mathbb{R}^d)$?

Now we prove the main result of this section.

(6.4) Theorem. *There exists a unique valuation $\mu : \mathcal{C}(\mathbb{R}^d) \longrightarrow \mathbb{R}$, called the Euler characteristic, such that $\mu([A]) = 1$ for every non-empty convex closed set $A \subset \mathbb{R}^d$.*

Proof. To show that μ must be unique, if it exists, is easy: let $f = \sum_{i=1}^{m} \alpha_i [A_i]$. Then we must have $\mu(f) = \sum_{i : A_i \neq \emptyset} \alpha_i$.

Let us prove that μ exists. First, we define μ on functions $f \in \mathcal{K}(\mathbb{R}^d)$.

We use induction on d. Suppose that $d = 0$. Then any function $f \in \mathcal{K}(\mathbb{R}^d)$ has the form $f = \alpha[0]$ for some $\alpha \in \mathbb{R}$ and we let $\mu(f) = \alpha$.

Suppose that $d > 0$.

For a point $x = (\xi_1, \ldots, \xi_d)$, let $l(x) = \xi_d$ be the last coordinate of x. For a $\tau \in \mathbb{R}$ let us consider the hyperplane

$$H_\tau = \big\{ x \in \mathbb{R}^d : l(x) = \tau \big\}.$$

The hyperplane H_τ can be identified with $\mathbb{R}^{d-1}$ and hence, by the induction hypothesis, there exists a valuation, say $\mu_\tau : \mathcal{K}(H_\tau) \longrightarrow \mathbb{R}$, which satisfies the required properties. For a function $f \in \mathcal{K}(\mathbb{R}^d)$, let f_τ be the restriction of f onto H_τ. Thus

$$\text{if} \quad f = \sum_{i=1}^{m} \alpha_i [A_i] \quad \text{then} \quad f_\tau = \sum_{i=1}^{m} \alpha_i [A_i \cap H_\tau]$$

and so $f_\tau \in \mathcal{K}(H_\tau)$ and we can define $\mu_\tau(f_\tau)$. Since $A_i \cap H_\tau$ are convex compact (possibly empty) sets we must have

$$\mu_\tau(f_\tau) = \sum_{i: A_i \cap H_\tau \neq \emptyset} \alpha_i.$$

Therefore, $\lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}(f_{\tau-\epsilon})$ is the sum of α_i such that $A_i \cap H_{\tau-\epsilon} \neq \emptyset$ for all sufficiently small $\epsilon > 0$. It follows then that

$$\mu_\tau(f_\tau) - \lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}(f_{\tau-\epsilon}) = \sum_{i \in I} \alpha_i, \quad \text{where} \quad I = \big\{ i : \min_{x \in A_i} l(x) = \tau \big\}.$$

In particular, $\lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}(f_{\tau-\epsilon}) = \mu_\tau(f_\tau)$ unless τ is the minimal value of the linear function $l(x)$ on some set A_i.

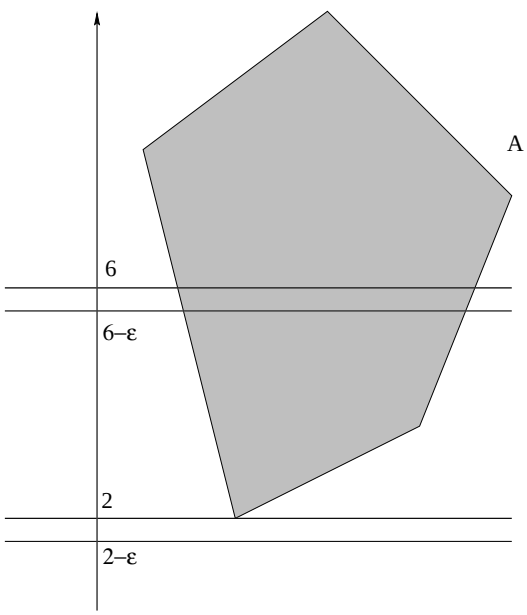


Figure 8

For example, for the function $f = [A]$ on Figure 8, we have $\lim_{\epsilon \longrightarrow +0} \mu_{6-\epsilon}(f_{6-\epsilon}) = \mu(f_6)$ but $\lim_{\epsilon \longrightarrow +0} \mu(f_{2-\epsilon}) \neq \mu(f_2)$.

Therefore, for a given function $f \in \mathcal{K}(\mathbb{R}^d)$ there are only finitely many τ's, where $\lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}(f_{\tau-\epsilon}) \neq \mu_\tau(f_\tau)$. Now we define

$$\mu(f) = \sum_{\tau \in \mathbb{R}} \Big(\mu_\tau(f_\tau) - \lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}(f_{\tau-\epsilon}) \Big).$$

As we noted, the sum contains only finitely many non-zero summands, so it is well-defined.

If $f, g \in \mathcal{K}(\mathbb{R}^d)$ are functions and $\alpha, \beta \in \mathbb{R}$ are numbers, then for every $\tau \in \mathbb{R}$ we have $(\alpha f + \beta g)_\tau = \alpha f_\tau + \beta g_\tau$. Since by the induction hypothesis μ_τ is a valuation and limit is a linear operation, we conclude that $\mu(\alpha f + \beta g) = \alpha\mu(f) + \beta\mu(g)$, so μ is a valuation. Furthermore, if $A \subset \mathbb{R}^d$ is a convex compact set, then

$$\mu_\tau([A]_\tau) - \lim_{\epsilon \longrightarrow +0} \mu_{\tau-\epsilon}([A]_{\tau-\epsilon}) = \begin{cases} 1 & \text{if } \min_{x \in A} l(x) = \tau \\ 0 & \text{otherwise.} \end{cases}$$

Since A is a non-empty convex compact set, there is unique minimal value of the linear function $l(x)$ on A. Therefore, $\mu([A]) = 1$.

Now we are ready to extend μ onto $\mathcal{C}(\mathbb{R})^d$. Let $B_r = \left\{x \in \mathbb{R}^d : \|x\| \leq r\right\}$ be the ball of radius r. For $f \in \mathcal{C}(\mathbb{R}^d)$ we define

$$\mu(f) = \lim_{r \longrightarrow +\infty} f \cdot [B].$$

It is seen that μ satisfies the required properties. □

Theorem 6.4 and its proof belongs to H. Hadwiger.

If $A \subset \mathbb{R}^d$ is a set such that $[A] \in \mathcal{C}(\mathbb{R}^d)$, we often write $\mu(A)$ instead of $\mu([A])$ and call it the Euler characteristic of the set A. In the course of the proof of Theorem 6.4, we proved the following important

(6.5) Lemma. *Let $A \subset \mathbb{R}^d$ be a set such that $[A] \in \mathcal{K}(\mathbb{R}^d)$. For $\tau \in \mathbb{R}$ let H_τ be the hyperplane consisting of the points $x = (\xi_1, \ldots, \xi_d)$ with $\xi_d = \tau$. Then $[A \cap H_\tau] \in \mathcal{K}(\mathbb{R}^d)$ and*

$$\mu(A) = \sum_{\tau \in \mathbb{R}} \Big(\mu(A \cap H_\tau) - \lim_{\epsilon \longrightarrow +0} \mu(A \cap H_{\tau-\epsilon})\Big).$$

□

Another important result is

(6.6) Corollary. *Let $A_1, \ldots, A_m \subset \mathbb{R}^d$ be sets such that $[A_i] \in \mathcal{K}(\mathbb{R}^d)$ for all $i = 1, \ldots, m$. Then $[A_1 \cup \ldots \cup A_m] \in \mathcal{K}(\mathbb{R}^d)$ and*

$$\mu(A_1 \cup \ldots \cup A_m) = \sum_{k=1}^{m} (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq m} \mu(A_{i_1} \cap \ldots \cap A_{i_k}).$$

In particular,

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2).$$

Proof. Follows by Lemma 6.2 and Theorem 6.4. □

PROBLEMS.

1. Let $A_1, A_2, A_3 \subset \mathbb{R}^d$ be closed convex sets such that $A_1 \cap A_2 \neq \emptyset$, $A_1 \cap A_3 \neq \emptyset$, $A_2 \cap A_3 \neq \emptyset$ and $A_1 \cup A_2 \cup A_3$ is a convex set. Prove that $A_1 \cap A_2 \cap A_3 \neq \emptyset$.

2. Let $A_1, \ldots, A_m \subset \mathbb{R}^d$ be closed convex sets such that $A_1 \cup \ldots \cup A_m$ is a convex set. Suppose that the intersection of every k sets $A_{i_1}, \ldots, A_{i_k}$ is non-empty. Prove that there are $k+1$ sets $A_{i_1}, \ldots, A_{i_{k+1}}$ whose intersection is non-empty.

3. Let $\Delta = \big\{(\xi_1, \ldots, \xi_d) : \xi_1 + \ldots + \xi_d = 1, \xi_i \geq 0 \text{ for } i = 1, \ldots, d\big\}$, $\Delta \subset \mathbb{R}^d$ be the standard simplex. Let $\Delta_i = \big\{x \in \Delta : \xi_i = 0\big\}$ be the i-th facet of Δ. Suppose that there are convex compact sets $K_1, \ldots, K_d \subset \mathbb{R}^d$, such that $\Delta \subset K_1 \cup \ldots \cup K_d$ and $K_i \cap \Delta_i = \emptyset$ for $i = 1, \ldots, d$. Prove that $K_1 \cap \ldots \cap K_d \neq \emptyset$.

Hint: Use induction on d and Problem 2.

4. Let $A_1, \ldots, A_m \subset \mathbb{R}^d$ be closed convex sets such that $A_1 \cap \ldots \cap A_m \neq \emptyset$. Prove that $\mu(A_1 \cup \ldots \cup A_m) = 1$.

5. Find the Euler characteristic of the "open square" $I_2 = \big\{(\xi_1, \xi_2) : 0 < \xi_1, \xi_2 < 1\big\}$ and the "open cube" $I_3 = \big\{(\xi_1, \xi_2) : 0 < \xi_1, \xi_2, \xi_3 < 1\big\}$.

6. Let $A_1, A_2, A_3, A_4 \subset \mathbb{R}^d$ be convex closed sets such that the union $A_1 \cup A_2 \cup A_3 \cup A_4$ is convex and all pairwise intersections $A_1 \cap A_2, A_1 \cap A_3, A_1 \cap A_4, A_2 \cap A_3, A_2 \cap A_4, A_3 \cap A_4$ are non-empty. Prove that at least 3 of the 4 intersections $A_1 \cap A_2 \cap A_3, A_1 \cap A_2 \cap A_4, A_1 \cap A_3 \cap A_4, A_2 \cap A_3 \cap A_4$ are non-empty and that if all the 4 intersections are non-empty, then the intersection $A_1 \cap A_2 \cap A_3 \cap A_4$ is non-empty. Construct an example, where exactly 3 of the 4 intersections $A_1 \cap A_2 \cap A_3, A_1 \cap A_2 \cap A_4, A_1 \cap A_3 \cap A_4, A_2 \cap A_3 \cap A_4$ are non-empty.

7. Application: Convex Sets and Linear Transformations

As an application of the Euler characteristic, we exhibit an interesting behavior of collections of compact convex sets under linear transformations (projections).

(7.1) Theorem. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. Then there exists a linear transformation $\mathcal{T} : \mathcal{K}(\mathbb{R}^n) \longrightarrow \mathcal{K}(\mathbb{R}^m)$, such that $\mathcal{T}([A]) = [T(A)]$ for any compact convex set $A \subset \mathbb{R}^n$.*

Proof. Clearly, if $A \subset \mathbb{R}^n$ is a convex compact set then $T(A) \subset \mathbb{R}^m$ is also a convex compact set. Let us define a function $G : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$, where

$$G(x,y) = \begin{cases} 1 & \text{if } T(x) = y \\ 0 & \text{if } T(x) \neq y. \end{cases}$$

Let $f \in \mathcal{K}(\mathbb{R}^n)$ be a function. We claim that for every $y \in \mathbb{R}^m$, the function $g_y(x) = G(x,y)f(x)$ belongs to the space $\mathcal{K}(\mathbb{R}^n)$. Indeed, if

$$f = \sum_{i=1}^{k} \alpha_i [A_i], \tag{7.1.1}$$

where $\alpha_i \in \mathbb{R}$ and $A_i \subset \mathbb{R}^n$ are convex compact sets, then

$$g_y = \sum_{i=1}^{k} \alpha_i [A_i \cap T^{-1}(y)], \tag{7.1.2}$$

where $T^{-1}(y)$ is the affine subspace that is the inverse image of y. Hence $\mu(g_y)$ is well-defined and we define $h = \mathcal{T}(f)$ by the formula $h(y) = \mu(g_y)$. We claim that $h \in \mathcal{K}(\mathbb{R}^m)$. Indeed, for f as in (7.1.1), the function g_y is given by (7.1.2) and

$$h(y) = \sum_{i \in I} \alpha_i, \quad \text{where} \quad I = \big\{i : A_i \cap T^{-1}(y) \neq \emptyset\big\}.$$

However, $A_i \cap T^{-1}(y) \neq \emptyset$ if and only if $y \in T(A_i)$, so

$$h = \sum_{i \in I} \alpha_i [T(A_i)]. \tag{7.1.3}$$

Therefore, $h = \mathcal{T}(f) \in \mathcal{K}(\mathbb{R}^m)$ and the transformation $\mathcal{T}$ is well-defined. We see that $\mathcal{T}$ is linear since for $f = \alpha_1 f_1 + \alpha_2 f_2$, we get $g_y(x) = \alpha_1 g_{1,y}(x) + \alpha_2 g_{2,y}(x)$, where $g_y(x) = G(y,x)f(x), g_{1,y} = G(y,x)f_1(x)$ and $g_{2,y} = G(y,x)f_2(x)$. Since μ is a linear functional (see Theorem 6.4), $h(y) = \alpha_1 h_1(y) + \alpha_2 h_2(y)$, where $h = \mathcal{T}(f)$, $h_1 = \mathcal{T}(f_1)$ and $h_2 = \mathcal{T}(f_2)$. It follows from (7.1.3) that $\mathcal{T}[A] = [T(A)]$. □

(7.2) Corollary. *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation, let $A_1, \ldots, A_k \subset \mathbb{R}^n$ be convex compact sets and let $\alpha_1, \ldots, \alpha_k$ be numbers such that*

$$\alpha_1[A_1] + \ldots + \alpha_k[A_k] = 0.$$

Then

$$\alpha_1[T(A_1)] + \ldots + \alpha_k[T(A_k)] = 0.$$

Proof. We apply the transformation $\mathcal{T}$ to both sides of the identity $\alpha_1[A_1] + \ldots + \alpha_k[A_k] = 0$. □

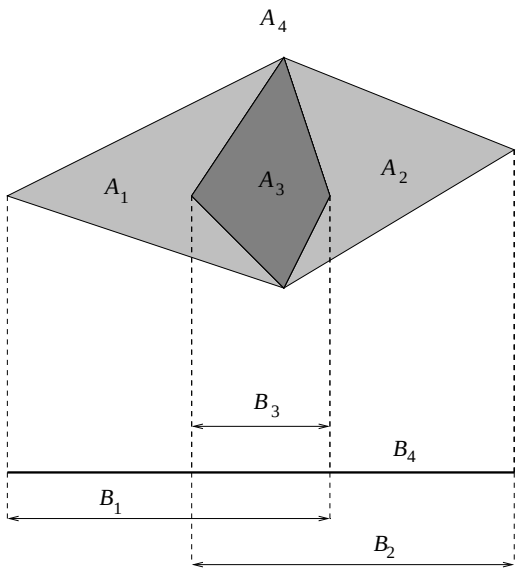


Figure 9

Figure 9 shows four convex sets A_1, A_2, A_3, A_4 such that $[A_4] = [A_1] + [A_2] - [A_3]$ and their projections B_1, B_2, B_3, B_4. We observe that $[B_4] = [B_1] + [B_2] - [B_3]$.

PROBLEMS.

1. For sets $A, B \subset \mathbb{R}^d$, we define their *Minkowski sum* $A + B = \{x + y : x \in A, y \in B\}$. Prove that the Minkowski sum of convex compact sets is a convex compact set and that there exists a commutative and associative operation $f \star g$ for functions $f, g \in \mathcal{K}(\mathbb{R}^d)$ such that $(\alpha_1 f_1 + \alpha_2 f_2) \star g = \alpha_1(f_1 \star g) + \alpha_2(f_2 \star g)$ for any $f_1, f_2, g \in \mathcal{K}(\mathbb{R}^d)$ and any $\alpha_1, \alpha_2 \in \mathbb{R}$ and such that $[A_1] \star [A_2] = [A_1 + A_2]$ for any convex compact sets $A_1, A_2 \subset \mathbb{R}^d$.

2. Construct an example of non-convex compact sets $A_i \subset \mathbb{R}^n$ and real numbers α_i such that $\sum_{i=1}^k \alpha_i [A_i] = 0$ but $\sum_{i=1}^k \alpha_i [T(A_i)] \neq 0$ for some linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

3. Construct an example of convex non-compact sets $A_i \subset \mathbb{R}^n$ and real numbers α_i such that $\sum_{i=1}^k \alpha_i [A_i] = 0$ but $\sum_{i=1}^k \alpha_i [T(A_i)] \neq 0$ for some linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

(7.3) Some interesting valuations. Intrinsic volumes. Let $\mathrm{vol}_d(A)$ be the usual volume of a convex compact set $A \subset \mathbb{R}^d$. The function vol_d satisfies a number of useful properties:

(7.3.1) Volume is (finitely) additive: If $A_1, \ldots, A_m \subset \mathbb{R}^d$ are convex compact sets and if $\alpha_1, \ldots, \alpha_m$ are numbers such that $\alpha_1 [A_1] + \ldots + \alpha_m [A_m] = 0$, then $\alpha_1 \mathrm{vol}_d(A_1) + \ldots + \alpha_m \mathrm{vol}_d(A_m) = 0$.

(7.3.2) Volume is invariant under *isometries* of $\mathbb{R}^d$, that is, orthogonal transformations and translations: $\mathrm{vol}_d\big(T(A)\big) = \mathrm{vol}_d(A)$ for any isometry $T : \mathbb{R}^d \longrightarrow \mathbb{R}^d$.

(7.3.3) Volume of a convex compact set $A \subset \mathbb{R}^d$ with non-empty interior is positive.

(7.3.4) Volume in $\mathbb{R}^d$ is homogeneous of degree d: $\mathrm{vol}_d(\alpha A) = \alpha^d \mathrm{vol}_d(A)$ for $\alpha \geq 0$, where $\alpha A = \{\alpha x : x \in A\}$.

It turns out that for every $k = 0, \ldots, d$, there exists a measure w_k on convex compact sets in $\mathbb{R}^d$, which satisfies properties (7.3.1)–(7.3.3) and which is homogeneous of degree k: $w_k(\alpha A) = \alpha^k w_k(A)$ for $\alpha > 0$. These measures are called *intrinsic volumes*. For $k = d$ we get the usual volume and for $k = 0$ we get the Euler characteristic.

To construct the intrinsic volumes, we observe that the volume can be extended to a valuation $\omega_d : \mathcal{K}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ such that $\omega_d([A]) = \mathrm{vol}_d(A)$ for any convex compact set A. Indeed, we define

$$\omega_d(f) = \int_{\mathbb{R}^d} f(x)\, dx,$$

where dx is the usual Lebesgue measure on $\mathbb{R}^d$. Properties of the integral imply that $\omega_d(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \omega_d(f_1) + \alpha_2 \omega_d(f_2)$, so ω_d is a valuation.

Let $L \subset \mathbb{R}^d$ be a k-dimensional subspace and let P_L be the orthogonal projection $P_L : \mathbb{R}^d \longrightarrow L$. Using Theorem 7.1, let us construct a linear transformation $\mathcal{P}_L : \mathcal{K}(\mathbb{R}^d) \longrightarrow \mathcal{K}(L)$ and hence a valuation $\omega_{k,L} : \mathcal{K}(L) \longrightarrow \mathbb{R}$ by letting $\omega_{k,L}(f) = \omega_k\big(\mathcal{P}_L(f)\big)$. Thus, for a convex compact set $A \subset \mathbb{R}^d$, the value of $\omega_{k,L}([A])$ is the volume of the orthogonal projection of A onto $L \subset \mathbb{R}^d$.

The functional $\omega_{k,L}[A]$ satisfies (7.3.1) and (7.3.3), it is homogeneous of degree k, but is not invariant under orthogonal transformations (although it is invariant under translations). To construct an invariant functional, we average $\omega_{k,L}$ over all k-dimensional subspaces $L \subset \mathbb{R}^d$. Let $G_k(\mathbb{R}^d)$ be the set of all k-dimensional subspaces $L \subset \mathbb{R}^d$. It is known that $G_k(\mathbb{R}^d)$ possesses a manifold structure (it is called the *Grassmannian*) and the rotationally invariant probability measure dL. Hence, for $f \in \mathcal{K}(\mathbb{R}^d)$ we let

$$\omega_k(f) = \int_{G_k(\mathbb{R}^d)} \omega_{k,L}(f)\, dL.$$

It is seen that $\omega_k : \mathcal{K}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ is a valuation. For a convex compact set $A \subset \mathbb{R}^d$ we define $w_k(A) := \omega_k([A])$.

Hence $w_k(A)$ is the average volume of projections of A onto k-dimensional subspaces in $\mathbb{R}^d$. The number $w_k(A)$ is called the k-th *intrinsic volume* of A. It satisfies properties (7.3.1)–(7.3.3) and it is homogeneous of degree k: $w_k(\alpha A) = \alpha^k w_k(A)$ for $\alpha \geq 0$. It is convenient to agree that $w_0(A) = \mu(A)$ and that $w_d(A) = \mathrm{vol}_d(A)$.

A good reference is D. A. Klain and G-C. Rota, *Introduction to Geometric Probability*, Cambridge University Press, 1997.

PROBLEMS.

1*. Let $A \subset \mathbb{R}^d$ be a compact convex set with non-empty interior. Prove that the *surface area* of A (*perimeter*, if $d = 2$) is equal $c_d w_{d-1}(A)$, where c_d is a constant depending on d alone. Find c_d.

2. Compute the intrinsic volumes of the unit ball $B = \big\{x \in \mathbb{R}^d : \|x\| \leq 1\big\}$.

Another interesting valuation.

3. Let us fix a vector $c \in \mathbb{R}^d$. For a non-empty convex compact set $A \subset \mathbb{R}^d$, let

$$h(A; c) = \max_{x \in A} \langle c, x \rangle$$

(when A is fixed, the function $h(A, c) : \mathbb{R}^d \longrightarrow \mathbb{R}$ is called the *support function* of A). Prove that there exists a valuation $\nu_c : \mathcal{K}(\mathbb{R}^d) \longrightarrow \mathbb{R}$ such that $\nu_c([A]) = h(A; c)$ for every non-empty convex compact set $A \subset \mathbb{R}^d$.

Hint: If $c \neq 0$, let

$$\nu_c(f) = \sum_{\alpha \in \mathbb{R}} \alpha \Big(\mu(f_\alpha) - \lim_{\epsilon \longrightarrow +0} \mu(f_{\alpha+\epsilon}) \Big),$$

where f_α is the restriction of f onto the hyperplane $H = \{x : \langle c, x\rangle = \alpha\}$.

4^*. Let $K_1, K_2 \subset \mathbb{R}^d$ be convex compact sets such that $K_1 \cup K_2$ is convex. Prove that $(K_1 \cup K_2) + (K_1 \cap K_2) = K_1 + K_2$, where "+" is the Minkowski sum (see Problem 1, Section 7.2).

Hint: Note that $[K_1 \cup K_2] + [K_1 \cap K_2] = [K_1] + [K_2]$ and use Problem 3 to conclude that $h(K_1 \cup K_2; c) + h(K_1 \cap K_2; c) = h(K_1; c) + h(K_2; c)$ for any $c \in \mathbb{R}^d$. Observe that $h(A + B; c) = h(A; c) + h(B; c)$.

8. The Isolation Theorem

We start with some linear algebra.

(8.1) Affine subspaces

Let V be a vector space and let $L \subset V$ be a subspace of V. For a vector $u \in V$ let

$$L + u = \{x + u : x \in L\}$$

be a translation of L. The set $A = L + u$ is called an *affine subspace* of V. The dimension of A is the dimension of L. In particular, if $\dim A = 1$, the set A is called a *straight line*. A straight line can be written in the parametric form $A = \{u + \alpha v : \alpha \in \mathbb{R}\}$, where $u, v \in V$ are vectors and $v \neq 0$. We say that $A = L + v$ is *parallel* to L. Another important example of an affine subspace is an affine hyperplane: Let $f : V \longrightarrow \mathbb{R}$ be a linear function (that is, $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$ for each $u, v \in V$ and each $\alpha, \beta \in \mathbb{R}$), which is not identically zero. The set

$$H = \{v \in V : f(v) = \alpha\}, \quad \text{where } \alpha \in \mathbb{R}$$

is called an *affine hyperplane*.

A linear combination

$$v = \alpha_1 v_1 + \ldots + \alpha_m v_m \quad \text{where} \quad \alpha_1 + \ldots + \alpha_m = 1$$

is called an *affine combination*. Similarly, points $v_1, \ldots, v_m \in V$ are said to be *affinely independent* if, whenever

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0 \quad \text{and} \quad \alpha_1 + \ldots + \alpha_m = 0$$

we must have $\alpha_1 = \ldots = \alpha_m = 0$.

Given a set $X \subset V$, the set $\mathrm{aff}(X)$ of all affine combinations of points from X is called the *affine hull* of X.

PROBLEMS.

1°. Prove that an affine combination of vectors from an affine subspace is a vector from the subspace.

2°. Prove that the intersection of affine subspaces is an affine subspace.

3°. Prove that an affine hyperplane is an affine subspace.

4°. Prove that an affine subspace is a subspace if and only if it contains 0.

5°. Let $A \subset V$ be an affine subspace of dimension n. Prove that the maximum number of affinely independent points in A is $n+1$.

6°. Let $L \subset V$ be a subspace and let $v, u \in V$ be vectors. Then $L + u = L + v$ if and only if $u - v \in L$.

7°. Let $X \subset V$ be a set of points in a vector space V. Prove that $\mathrm{aff}(X)$ is the smallest affine subspace containing X.

(8.2) Quotients, projections and codimension

Let V be a vector space and let $L \subset V$ be a subspace. We can form the *quotient space* V/L as follows: the points of V/L are affine subspaces of the type $L + v$, where $v \in V$. The addition in V/L is defined as follows: $A_1 + A_2 = A_3$ provided $A_1 = L + v_1$, $A_2 = L + v_2$, $A_3 = L + v_3$ and $v_1 + v_2 = v_3$ for some v_1, v_2 and v_3. The scalar multiplication in V/L is defined as follows: $\alpha A_1 = A_2$ provided $A_1 = L + v_1$, $A_2 = L + v_2$ and $v_2 = \alpha v_1$ for some $v_1, v_2 \in V$. Thus L itself is the 0 of the quotient V/L. The dimension of V/L is called the *codimension* of L (denoted $\operatorname{codim} L$). If $A = L + u$ is an affine subspace, the codimension of A is defined to be the codimension of L. There is a linear transformation, called *projection* $pr : V \longrightarrow V/L$, where $pr(v) = L + v$.

PROBLEMS.

1°. Prove that the addition and scalar multiplication in V/L are well-defined (do not depend on the particular choice of v_1 and v_2).

Addition: let $L \subset V$ be a subspace. Suppose there are two vectors u_1, v_1 such that $L + u_1 = L + v_1 = A_1$ and two vectors u_2, v_2 such that $L + u_2 = L + v_2 = A_2$. Let $u_3 = u_1 + u_2$ and $v_3 = v_1 + v_2$. Then $L + v_3 = L + u_3 = A_3$.

Multiplication: let $L \subset V$ be a subspace. Suppose there are two vectors u_1, v_1 such that $L + u_1 = L + v_1 = A_1$. For $\alpha \in \mathbb{R}$ let $v_2 = \alpha v_1$ and $u_2 = \alpha u_1$. Then $L + u_2 = L + v_2 = A_2$.

2°. Prove that the affine hyperplanes are exactly the affine subspaces of codimension 1.

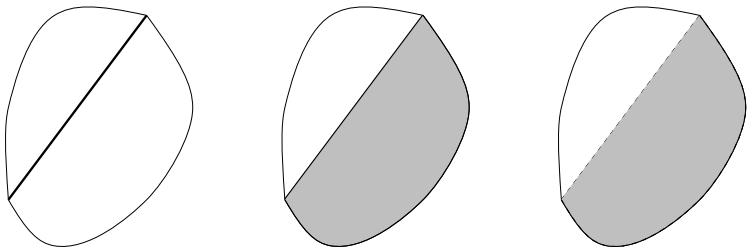
3°. Prove that the projection $V \longrightarrow V/L$ is indeed a linear transformation, that its image is the whole space V/L, and that its kernel is L.

4°. Let $L \subset \mathbb{R}^d$ be a subspace. Prove that $\dim L + \operatorname{codim} L = d$.

(8.3) Halfspaces

Let V be a vector space and let $H \subset V$ be an affine hyperplane. Then the complement of H in V is the union of two convex sets, called *open halfspaces*: $V \setminus H = H_+ \cup H_-$. Indeed, suppose that $H = \{x \in \mathbb{R}^d : f(x) = \alpha\}$, where $f : V \longrightarrow \mathbb{R}$ is a linear function and $\alpha \in \mathbb{R}$ is a number. We let $H_+ = \{x \in \mathbb{R}^d : f(x) > \alpha\}$ and $H_- = \{x \in \mathbb{R}^d : f(x) < \alpha\}$. Of course, if we choose a different equation for H (say, $-f(x) = -\alpha$) then H_+ and H_- interchange. The

sets $\overline{H_+} = H \cup H_+$ and $\overline{H_-} = H \cup H_-$ are called *closed subspaces*. We can write $\overline{H_+} = \big\{x \in \mathbb{R}^d : f(x) \geq \alpha\big\}$ and $\overline{H_-} = \big\{x \in \mathbb{R}^d : f(x) \leq \alpha\big\}$.



Hyperplane Closed halfspace Open halfspace

Figure 10

PROBLEMS.

1°. Prove that open halfspaces and closed halfspaces are convex.

2°. Prove that an open halfspace of $\mathbb{R}^d$ is an open subset of $\mathbb{R}^d$ and that a closed halfspace of $\mathbb{R}^d$ is a closed subset of $\mathbb{R}^d$.

(8.4) Definitions. Let V be a vector space, let $A \subset V$ be a set and let $H \subset V$ be a hyperplane. We say that H *isolates* A if A is contained in one of the closed subspaces $\overline{H_-}$ or $\overline{H_+}$. We say that H *strictly isolates* A if A is contained in one of the open halfspaces H_- or H_+.

Let V be a vector space, let $A, B \subset V$ be sets and let $H \subset V$ be a hyperplane. We say that H *separates* A and B if A is contained in one closed halfspace and B is contained in the other. We say that H *strictly separates* A and B if A is contained in one open halfspace and B is contained in the other open halfspace.

PROBLEM.

1°. Prove that sets $A, B \subset V$ can be separated, resp. strictly separated, by a hyperplane if and only if there is a linear function $f : V \longrightarrow \mathbb{R}^d$ and a number $\alpha \in \mathbb{R}$ such that $f(x) \leq \alpha \leq f(y)$, resp. $f(x) < \alpha < f(y)$ for any $x \in A$ and any $y \in B$.

(8.5) Definition. Let V be a vector space and let $A \subset V$ be a set. The set A is called *algebraically open* if the intersection of A with every straight line in V is a union of open intervals (possibly empty). In particular, if A is convex and $L = \big\{v + \tau u : \tau \in \mathbb{R}\big\}$ is a straight line in V, where $v, u \in V$, then

$$A \cap L = \big\{v + \tau u : \alpha < \tau < \beta\big\}, \quad \text{where}$$

either $-\infty < \alpha < \beta < +\infty$ (the intersection is a non-empty open interval);
or $\alpha = -\infty < \beta < +\infty$ (the intersection is an open ray);

or $-\infty < \alpha < \beta = +\infty$ (the intersection is an open ray);
or $-\infty = \alpha < \beta = +\infty$ (the intersection is the whole straight line);
or $\alpha \geq \beta$ (the intersection is empty).

PROBLEMS.

1°. Prove that open sets in $\mathbb{R}^d$ are algebraically open.

2. Construct an example of an algebraically open set $A \subset \mathbb{R}^2$ which is not open.

3. Let $A \subset \mathbb{R}^d$ be a convex set. Prove that it is open if and only if it is algebraically open.

4. Prove that if an algebraically open set is isolated by an affine hyperplane, it is strictly isolated by the hyperplane.

5. Let V be a vector space, let $A, B \subset V$ be algebraically open sets. Prove that if A and B are separated by an affine hyperplane H then A and B are strictly separated by H.

6°. Prove that the union of algebraically open sets is algebraically open and that the intersection of a finite family of algebraically open sets is algebraically open.

7. Let V and W be vector spaces, let $T : V \longrightarrow W$ be a linear transformation such that $\mathrm{im}(T) = W$. Let $A \subset V$ be an algebraically open set in V. Then the image $T(A)$ is algebraically open in W.

8°. Let $A \subset V$ be an algebraically open set and let $L \subset V$ be a subspace. Prove that $A \cap L$ is algebraically open as a set in L.

Finally, the main result of this section.

(8.6) The Isolation Theorem. *Let V be a vector space, let $A \subset V$ be an algebraically open convex set and let $u \notin A$ be a point. Then there exists an affine hyperplane H which contains u and strictly isolates A.*

Proof. Without loss of generality we may assume that $u = 0$ is the origin.

First, we prove the result in the case of $V = \mathbb{R}^2$. Let $S = \{x \in \mathbb{R}^2 : \|x\| = 1\}$ be the circle centered at the origin of radius 1. Let us project radially A into S: $v \longrightarrow v/\|v\|$. Since A is convex, it is connected, and, therefore, the image of this projection is a connected arc Γ of S. Furthermore, since A is algebraically open, Γ must be an open arc

$$\Gamma = \Big\{(\cos\phi,\ \sin\phi) : \alpha < \phi < \beta\Big\}$$

of the circle S. Indeed, let $x \in \Gamma$ be a point. Then $x = v/\|v\|$ for some $v \in A$, so we can choose a straight line L through v parallel to the tangent line to S at x. Then the intersection $L \cap A$ will be an open interval containing v, so the radial projection of A will contain an open arc containing x. Next, we observe that the length of Γ can not be greater then π, because otherwise Γ would have contained two antipodal points x and $-x$ and 0 would have been in A, since A is convex. Now, let v be an

endpoint of Γ (which is not in Γ, since Γ is open). The straight line through 0 and v is the desired hyperplane, containing 0 and strictly isolating A.

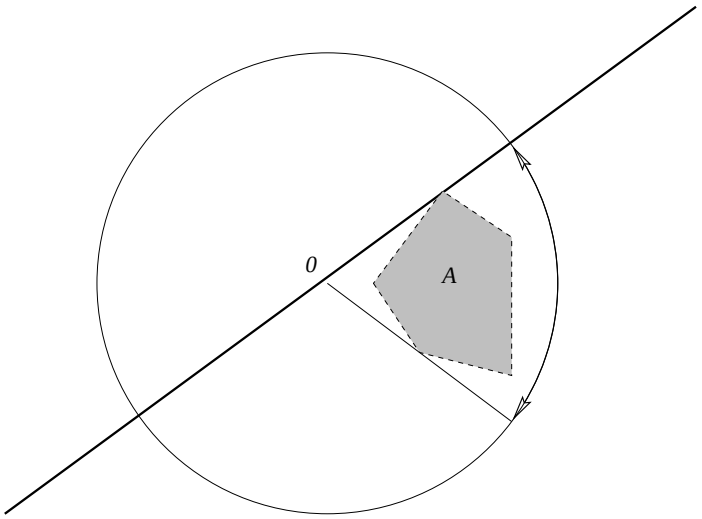


Figure 11

Next, suppose that $\dim V \geq 2$. We prove, that there is a straight line L, such that $0 \in L$ and $L \cap A = \emptyset$. To prove this, let us consider any 2-dimensional plane P containing 0. The intersection $B = P \cap A$ is a convex algebraically open subset of P (possibly empty, see Problem 8 of Section 8.5) and as we proved, there is a line $L \subset P$ such that $0 \in L$ and $L \cap B = \emptyset$. Then L is the desired straight line.

Now, we prove the theorem. Let $H \subset V$ be the largest affine subspace such that $0 \in H$ and $H \cap A = \emptyset$. By "largest", we mean a subspace which has these properties and is not contained in a larger subspace with the same properties. If V is finite-dimensional, we could chose H to be a subspace of the largest possible dimension such that $0 \in H$ and $H \cap A = \emptyset$. If V is arbitrary, the existence of such an H is ensured by Zorn's Lemma. We claim that H is a hyperplane. To prove this, consider the quotient V/H and let $pr : V \longrightarrow V/H$ be the projection. If H is not a hyperplane, then $\dim V/H \geq 2$ and $pr(A)$ is an algebraically open subset in V/H (see Problem 7, Section 8.5). Then, as we proved, there is a straight line $L \subset V/H$ such that $0 \in L$ and $L \cap pr(A) = \emptyset$. Then the preimage $G = pr^{-1}(L) = \big\{x : pr(x) \in L\big\}$ is a subspace in V, such that $0 \in G$, $G \cap A = \emptyset$, $H \subset G$ and G is strictly larger than H. This contradiction shows that H is a hyperplane. $\square$

PROBLEMS.

1°. Construct an example of a non-convex algebraically open set $A \subset \mathbb{R}^2$ such that $0 \notin A$ and there isn't an affine hyperplane H such that $0 \in H$ and H isolates A.

2. Let $V = \mathbb{R}_\infty$ be the vector space of all infinite sequences $x = (\xi_1, \dots)$ of real numbers such that all but finitely many terms ξ_i are zero. One can think of $\mathbb{R}_\infty$ as

of the space of all univariate polynomials with real coefficients. Let $A \subset V$ be the set of all such sequences x where the last non-zero term is strictly positive. Prove that $0 \notin A$, that A is convex, that A is not algebraically open and that there isn't an affine hyperplane H such that $0 \in H$ and H isolates A.

3. Prove the following generalization of Theorem 8.6: Let V be a vector space, let $A \subset V$ be an algebraically open convex set, and let $L \subset V$ be an affine subset such that $L \cap A = \emptyset$. Then there exists an affine hyperplane H containing L that strictly isolates A.

9. Convex Sets in Euclidean Space

In this Section, we investigate corollaries of the Isolation Theorem for convex sets in Euclidean space.

(9.1) Definition. Let $A \subset \mathbb{R}^d$ be a set. A point $u \in A$ is called an *interior* point of A if there exists an $\epsilon > 0$ such that the (open) ball $B(u, \epsilon) = \{x : \|x - u\| < \epsilon\}$ centered at u and of the radius ϵ is contained in A: $B(u, \epsilon) \subset A$. The set of all interior points of A is called the *interior* of A and denoted $\text{int}(A)$. The set of all non-interior points of A is called the *boundary* of A and denoted ∂A.

(9.2) Lemma. *Let $A \subset \mathbb{R}^d$ be a convex set and let $u_0 \in \text{int}(A)$ be an interior point of A. Then, for any point $u_1 \in A$ and any $0 \leq \alpha < 1$, the point $u_\alpha = (1-\alpha)u_0 + \alpha u_1$ is an interior point of A.*

Proof. Let $B(u_0, \epsilon) \subset A$ be a ball centered at u_0 and contained in A. Then elementary geometry shows that $B\big(u_\alpha, (1-\alpha)\epsilon\big) \subset A$, see Figure 12. □

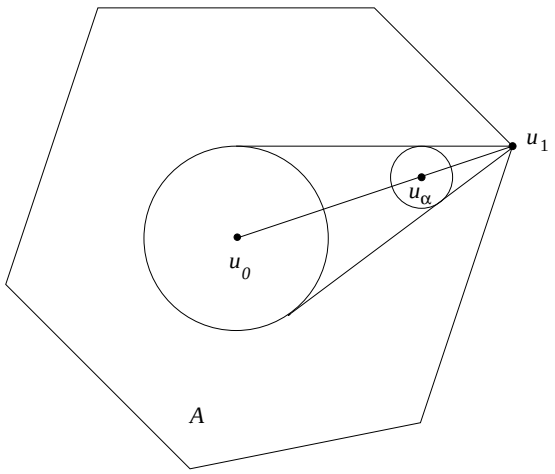


Figure 12

(9.3) Corollary. *Let $A \subset \mathbb{R}^d$ be a convex set. Then $\text{int}(A)$ is a convex set.*

Proof. Let $u, v \in \text{int}(A)$ be points and let $w = \alpha u + (1-\alpha)v$ for $0 \leq \alpha \leq 1$. If $\alpha < 1$ we apply Lemma 9.2 with $u_0 = v, u_1 = u$ and $w = u_\alpha$ to show that $w \in \text{int}(A)$. If $\alpha = 1$ then $w = u \in \text{int}(A)$. □

PROBLEMS.

1. Let $v_1, \ldots, v_{d+1} \in \mathbb{R}^d$ be affinely independent points in $\mathbb{R}^d$. The polytope $\Delta = \operatorname{conv}\big(v_1, \ldots, v_{d+1}\big)$ is called a d-dimensional simplex. Prove that Δ has a non-empty interior.

Hint: Let $u = (v_1 + \ldots + v_{d+1})/(d+1) \in \Delta$. We claim that for a sufficiently small $\epsilon > 0$ we have $B(u, \epsilon) \subset \Delta$. Indeed, the matrix of the system of $(d+1)$ linear equations in $(d+1)$ variables

$$\gamma_1 v_1 + \ldots + \gamma_{d+1} v_{d+1} = w, \quad \gamma_1 + \ldots + \gamma_{d+1} = 1$$

is non-degenerate. Therefore, for each $w \in \mathbb{R}^d$, there is a unique solution $\gamma_1, \ldots, \gamma_{d+1}$ and the solution depends on w continuously. If $w = u$ then $\gamma_1 = \ldots = \gamma_{d+1} = 1/(d+1) > 0$. Therefore, if w is sufficiently close to u, all the γ's are non-negative and $w \in \Delta$.

2. Let $A \subset \mathbb{R}^d$ be a convex set, such that $\operatorname{int}(A) \neq \emptyset$ and let $H \subset \mathbb{R}^d$ be a hyperplane isolating $\operatorname{int}(A)$. Prove that H isolates A.

We want to show that if a non-empty convex set in Euclidean space has empty interior, then we can pass to a smaller ambient space, where the set acquires an interior. This property makes the finite-dimensional situation so radically different from the infinite-dimensional case.

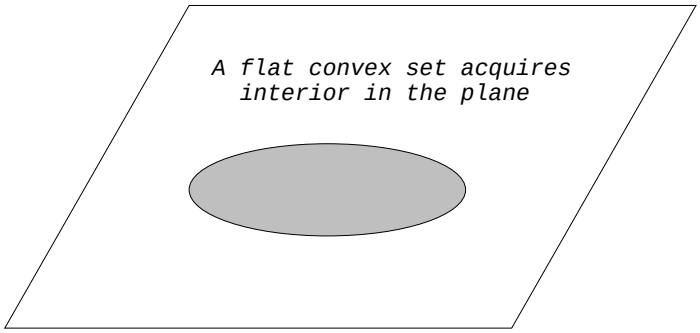


Figure 13

(9.4) Theorem. *Let $A \subset \mathbb{R}^d$ be a convex set. If $\operatorname{int} A = \emptyset$, then there exists an affine subspace $L \subset \mathbb{R}^d$ such that $A \subset L$ and $\dim L < d$.*

Proof. First, we claim there are no $d+1$ affinely independent points $v_1, \ldots, v_{d+1} \in A$. For if there were such points, then $\Delta = \operatorname{conv}\big(v_1, \ldots, v_{d+1}\big) \subset A$ and Problem 1, Section 9.3 would have implied that Δ contains an interior point. Let $k < d+1$ be the maximum number of affinely independent points in A and let $v_1, \ldots, v_k$ be such points. Then, for each point $v \in A$ there is a solution to the system

$$\gamma_1 v_1 + \ldots + \gamma_k v_k + \gamma v = 0$$
$$\gamma_1 + \ldots + \gamma_k + \gamma = 0$$

such that $\gamma \neq 0$. Then $v \in A$ can be expressed as an affine combination of $v_1, \ldots, v_k$,

$$v = \sum_{i=1}^{k} (-\gamma_i/\gamma) v_k.$$

Therefore, A is contained in the affine subspace L that is the affine hull of $v_1, \ldots, v_k$. So, $\dim L = k - 1$. □

(9.5) Definition. The *dimension* of a convex set $A \subset \mathbb{R}^d$ is the dimension of the smallest affine subspace that contains A. The dimension of the empty set is -1.

PROBLEM.

1. Let $A \subset \mathbb{R}_\infty$ be the set of Problem 2, Section 8.6. Prove that A does not contain any non-empty algebraically open subset and that A is not contained in any affine hyperplane in $\mathbb{R}_\infty$.

(9.6) Definitions. Let $K \subset \mathbb{R}^d$ be a closed convex set. A (possibly empty) set $F \subset K$ is called a *face* of K if there exists an affine hyperplane H which isolates K and such that $F = K \cap H$. If F is a point, then F is called an *exposed point* of K. A non-empty face $F \neq K$ is called a *proper* face of K.

PROBLEMS.

1°. Prove that a face is a closed convex set and that a face of a compact convex set is a compact convex set.

2. Find the faces of the unit ball $B = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$.

3. Describe the faces of the unit d-dimensional cube $I = \Big\{x = (\xi_1, \ldots, \xi_d) : 0 \leq \xi_k \leq 1 \text{ for } k = 1, \ldots, d\Big\}$.

4. Describe the faces of a d-dimensional simplex $\Delta = \operatorname{conv}\big(v_1, \ldots, v_{d+1}\big)$, where $v_1, \ldots, v_{d+1}$ are affinely independent points in $\mathbb{R}^d$.

5. Let $K \subset \mathbb{R}^d$ be a convex compact set. Prove that the intersection of any two faces of K is a face of K.

6. Construct an example of a convex compact set $K \subset \mathbb{R}^2$, a face F of K and a face G of F such that G is not a face of K.

7. Prove that every non-empty convex compact set in $\mathbb{R}^d$ has an exposed point.

8. Construct a convex compact set $A \subset \mathbb{R}^2$ whose set of exposed points is not compact.

9* (Straszewicz' Theorem). Prove that every convex compact set $A \subset \mathbb{R}^d$ is the closure of the convex hull of the set of its exposed points.

(9.7) Theorem. *Let $K \subset \mathbb{R}^d$ be a closed convex set with a non-empty interior and let $u \in \partial K$ be a point. Then there exists an affine hyperplane H, called a support hyperplane at u, such that $u \in H$ and H isolates K.*

Proof. By Corollary 9.3, $\operatorname{int}(K)$ is a non-empty convex open set. Therefore, $\operatorname{int}(K)$ is a convex, algebraically open set, such that $u \notin \operatorname{int}(K)$. Therefore, by Theorem

8.6, there is an affine hyperplane H containing u and isolating $\mathrm{int}(K)$. Then by Problem 2, Section 9.3, H isolates K, so H is a support hyperplane at u. $\square$

PROBLEM.

1°. Construct an example of a convex closed set $K \subset \mathbb{R}^2$ with a non-empty interior and a point $u \in \partial K$ such that a support hyperplane of K at u is not unique.

2°. Let $B = \left\{x \in \mathbb{R}^d : \|x\| \leq 1\right\}$ be the unit ball and let $u \in \partial B$ be a point. Find the support hyperplane to B at u.

(9.8) Corollary. *Let $K \subset \mathbb{R}^d$ be a closed convex set with a non-empty interior and let $u \in \partial K$ be a point. Then there is a proper face F of K such that $u \in F$.*

Proof. Let H be a support hyperplane of K at u. Let $F = H \cap K$. $\square$

Finally, we prove a version of the isolation theorem for convex sets in $\mathbb{R}^d$.

(9.9) Theorem. *Let $A \subset \mathbb{R}^d$ be a non-empty convex set and let $u \notin A$ be a point. Then there is an affine hyperplane $H \subset \mathbb{R}^d$ such that $u \in H$ and H isolates A.*

Proof. Let us choose the minimal affine subspace $L \subset \mathbb{R}^d$ such that $A \subset L$. Theorem 9.4 implies that A has a non-empty interior as a subset of L. If $u \notin L$, we can choose H parallel to L. Thus, restricting ourselves to L, we see that $\mathrm{int}(A) \neq \emptyset$ (in L) and that $u \in L$. Then, by Theorem 9.4, there is an affine hyperplane $\hat{H}$ in L, such that $u \in \hat{H}$ and $\hat{H}$ isolates $\mathrm{int}(A)$. Then, by Problem 2, Section 9.3, $\hat{H}$ isolates A. Then we choose any hyperplane H such that $H \cap L = \hat{H}$. $\square$

PROBLEM.

1. Let $A \subset \mathbb{R}^d$ be a convex set and let $L \subset \mathbb{R}^d$ be an affine subspace such that $L \cap A = \emptyset$. Prove that there exists an affine hyperplane H such that $L \subset H$ and H isolates A.

10. Extreme Points. The Krein-Milman Theorem in Euclidean Space

First, a very important definition.

(10.1) Extreme points. Let V be a vector space and let $A \subset V$ be a set. A point $a \in A$ is called an *extreme* point of A provided for any two points $b, c \in A$ such that $(b + c)/2 = a$, one must have $b = c = a$. The set of all extreme points of A is denoted $\mathrm{ex}(A)$.

A simple, yet important theorem.

(10.2) Theorem. *Let V be a vector space, let $A \subset V$ be a non-empty set, and let $f : V \longrightarrow \mathbb{R}$ be a linear function.*

(1) *Suppose that f attains its unique maximum on A at a point $u \in A$, that is $f(u) > f(v)$ for any $v \neq u, v \in A$. Then u is an extreme point of A;*
(2) *Suppose that f attains its maximum α on A and suppose that $B = \left\{x \in A : f(x) = \alpha\right\}$ is the set where the maximum is attained. Let u be an extreme point of B. Then u is an extreme point of A.*

Proof. Let us prove the first part. If $u = (a+b)/2$ then $f(u) = (f(a) + f(b))/2$, where $f(a) \le f(u)$ and $f(b) \le f(u)$. Therefore, $f(a) = f(b) = f(u)$ and we must have $a = b = u$, because the maximum is unique. Let us prove the second part. Suppose that $u = (a+b)/2$ for $a, b \in A$. Then $\alpha = f(u) = (f(a) + f(b))/2$ and $f(a), f(b) \le \alpha$. Thus we must have $f(a) = f(b) = \alpha$, so $a, b \in B$. Then $a = b = u$ since u is an extreme point of B. $\square$

PROBLEMS.

1°. Let $K \subset \mathbb{R}^d$ be a convex closed set and let $F \subset K$ be a face. Prove that if $u \in F$ is an extreme point of F then u is an extreme point of K.

2. Let $K \subset \mathbb{R}^d$ be a compact convex set and let $u \in K$ be a point such that $\|u\| \ge \|v\|$ for each $v \in K$. Prove that u is an extreme point of K.

The following remarkable result was obtained by M.G. Krein and D.P. Milman in 1940. Here we present its finite-dimensional version, an infinite-dimensional extension will be obtained later.

(10.3) Krein-Milman Theorem. *Let $K \subset \mathbb{R}^d$ be a convex compact set. Then K is the convex hull of the set of its extreme points: $K = \operatorname{conv}\big(\operatorname{ex}(K)\big)$.*

Proof. We proceed by induction on the dimension d. If $d = 0$ then K is a point and the result follows. Suppose that $d > 0$. Without loss of generality we may suppose that $\operatorname{int}(K) \ne \emptyset$ (otherwise, K lies in an affine subspace of a smaller dimension and the result follows by the induction conjecture). It suffices to prove that every point $u \in K$ can be represented as a convex combination of extreme points of K. If $u \in \partial K$, then, by Corollary 9.8, there exists a face F of K such that $u \in F$. Then F lies in an affine subspace of a smaller dimension, and by the induction conjecture $u \in \operatorname{conv}\big(\operatorname{ex}(F)\big)$, so the result follows since $\operatorname{ex}(F) \subset \operatorname{ex}(K)$ (see Problem 1, Section 10.2). Suppose that $u \in \operatorname{int}(K)$. Let us draw a straight line L trough u. The intersection $L \cap K$ is an interval $[a, b]$, where $a, b \in \partial K$ and u is an interior point of $[a, b]$. As we already proved, $a, b \in \operatorname{conv}\big(\operatorname{ex}(K)\big)$. Since u is a convex combination of a and b, the proof follows. $\square$

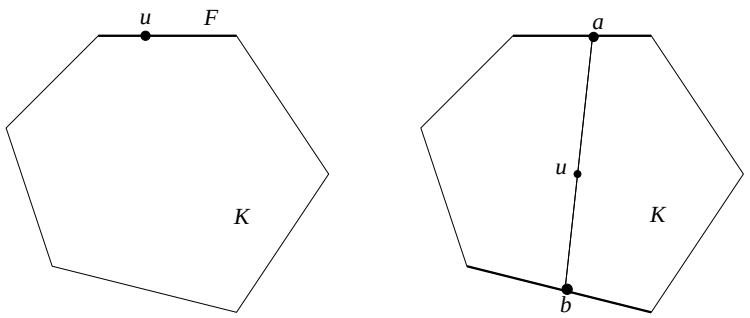


Figure 14

PROBLEMS.

1. Prove that the set of extreme points of a closed convex set in $\mathbb{R}^2$ is closed.

2. Construct an example of a compact convex set $K \subset \mathbb{R}^3$ such that $\mathrm{ex}(K)$ is not closed.

3. Let $A \subset \mathbb{R}^d$ be a set. Prove that u is an extreme point of $\mathrm{conv}(A)$ if and only if $u \in A$ and $u \notin \mathrm{conv}\big(A \setminus \{u\}\big)$.

4. Construct an example of a convex compact set $K \subset \mathbb{R}^2$ and a point $u \in K$ such that u is an extreme point of K, but not an exposed point of K.

5°. Prove that exposed point is an extreme point.

6*. Let $A \subset \mathbb{R}^d$ be a closed convex set. Prove that each extreme point of A is a limit of exposed points of A.

(10.4) Corollary. *Let $K \subset \mathbb{R}^d$ be a convex compact set and let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function. Then there exists an extreme point u of K such that $f(u) \geq f(x)$ for any $x \in K$.*

Proof. Clearly, f attains its maximum value, say, α on K. Let $F = \big\{x \in K : f(x) = \alpha\big\}$ be the corresponding face of K. Then $\mathrm{ex}(F) \neq \emptyset$ and any $u \in \mathrm{ex}(F)$ is an extreme point of K. □

11. Extreme Points of Polyhedra

(11.1) Definition. The intersection of finitely many closed halfspaces in $\mathbb{R}^d$ is called a *polyhedron*. An extreme point of a polyhedron is called a *vertex*.

(11.2) Theorem. *Let $P \subset \mathbb{R}^d$ be a polyhedron*

$$P = \Big\{x \in \mathbb{R}^d : \langle c_i, x\rangle \leq \beta_i : \ i = 1, \ldots, m\Big\},$$

where $c_i \in \mathbb{R}^d$ and $\beta_i \in \mathbb{R}$ for $i = 1, \ldots, m$.

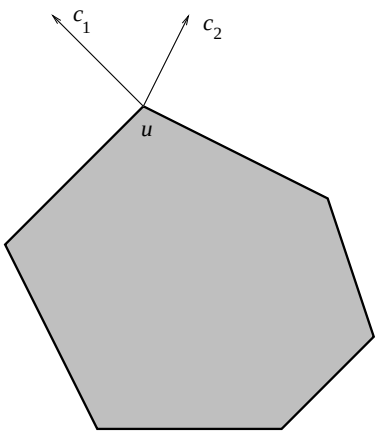
For $u \in P$ let

$$I(u) = \big\{i : \langle c_i, u\rangle = \beta_i\big\}$$

be the set of the inequalities that are "active" on u. Then u is a vertex of P if and only if the set of vectors $\{c_i : i \in I(u)\}$ spans the vector space $\mathbb{R}^d$. In particular, if u is a vertex of P, the set $I(u)$ contains at least d indices: $|I(u)| \geq d$.

Proof. Suppose that the vectors $c_i : i \in I(u)$ do not span $\mathbb{R}^d$. Then there is a non-zero $y \in \mathbb{R}^d$ such that $\langle y, c_i\rangle = 0$ for all $i \in I(u)$. We note that $\langle c_i, u\rangle < \beta_i$ for $i \notin I(u)$. For $\epsilon > 0$ let $u_+(\epsilon) = u + \epsilon y$ and let $u_-(\epsilon) = u - \epsilon y$. Then $u = (u_+(\epsilon) + u_-(\epsilon))/2$, $u_+(\epsilon) \neq u_-(\epsilon)$ and for sufficiently small $\epsilon > 0$ the points $u_-(\epsilon)$ and $u_+(\epsilon)$ belong to the polyhedron P. Hence u is not an extreme point of P.

Suppose now that $u \in P$ and the vectors $c_i : i \in I(u)$ span $\mathbb{R}^d$. Suppose that $u = (v+w)/2$ for $v, w \in P$. Then $\langle c_i, v\rangle \leq \beta_i$ and $\langle c_i, w\rangle \leq \beta_i$. Since $\langle c_i, u\rangle = \beta_i$ for $i \in I(u)$, we must have $\langle c_i, v\rangle = \langle c_i, w\rangle = \beta_i$ for $i \in I(u)$. Since vectors $c_i : i \in I(u)$ span $\mathbb{R}^d$, the system of linear equations must have a unique solution. Therefore, $v = w = u$ and u is an extreme point. □



A polyhedron and its vertex

Figure 15

PROBLEM.

1. Prove that every vertex of a polyhedron is an exposed point.

(11.3) Corollary. *Bounded polyhedron is a polytope, that is the convex hull of finitely many points.*

Proof. Every set of linear equations $\langle c_i, x\rangle = \beta_i : i \in I$, where $\{c_i : i \in I\}$ span $\mathbb{R}^d$, has at most one solution. Therefore, by Theorem 11.2, the number of vertices of a polyhedron in $\mathbb{R}^d$, defined by a set of m inequalities, does not exceed $\binom{m}{d}$, and hence is finite. By the Krein-Milman Theorem (Theorem 10.3), P is the convex hull of the set of its extreme points and the proof follows. □

PROBLEMS.

1. Prove that a polyhedron has finitely many faces.
2°. Prove that a face of a polyhedron is a polyhedron.
3. Prove that polytopes have finitely many faces.

Effect of "unrealistic solutions" in linear programming problems

Let $P \subset \mathbb{R}^d$ be a polyhedron defined by a system of m linear inequalities. Suppose we want to solve a *linear programming problem*

Find $\max\langle c_0, x\rangle$
Subject to: $x \in P$,

where $c_0 \in \mathbb{R}^d$ is the vector of the objective function. If the point $u \in P$, where the maximum is attained, is unique, then by Part 1, Theorem 10.2, u must be a vertex of P. Theorem 11.2 then implies that at least d of the m inequality constraints turn equalities on u. This sometimes is not at all desired.

(11.4) Example. The Diet Problem. Suppose we have m different food ingredients, the unit price of the i-th ingredient is γ_i, $i = 1, \dots, m$. We want the diet to be balanced with respect to k nutrients. Suppose that α_{ij} is the content of the j-th nutrients in the i-th ingredient. Let $\xi_i : i = 1, \dots, m$ be the quantity of the i-th ingredient in the diet and let $\beta_j : j = 1, \dots, k$ be the target quantity of the j-th nutrient.

Trying to find the cheapest balanced diet, we come to a *linear programming* problem:

$$\text{Find } \min \sum_{i=1}^{m} \gamma_i \xi_i$$

$$\text{Subject to: } \sum_{i=1}^{m} \alpha_{ij} \xi_i = \beta_j \text{ for } j = 1, \dots k;$$

$$\xi_i \geq 0 \text{ for } i = 1, \dots, m.$$

Let $P \subset \mathbb{R}^d$ be the polyhedron of all feasible diets $x = (\xi_1, \dots, \xi_d)$. Clearly, P lies in the affine subspace defined by the k balance constraints $\sum_{i=1}^{m} \alpha_{ij} \xi_i = \beta_j$. The dimension of the subspace, if it is non-empty, is at least $d - k$. If the optimal diet $x = (\xi_1, \dots, \xi_d)$ is unique, it has to be a vertex of P, so at least $d - k$ of the coordinates $\xi_1, \dots, \xi_d$ are zero. This, in turn, means that the optimal diet would consist of at most k ingredients. For example, if we balancing the diet by the content of 5 substances, we should expect to get the menu consisting of 5 or fewer ingredients only.

PROBLEM.

1°. In Section 1.9 of *Linear Algebra and its Applications* (second addition) by D.C. Lay, Addison-Wesley, 1997, the author describes the Cambridge diet. In particular, the last line of the penultimate paragraph on p. 87 reads: "In fact, the manufacturer of the Cambridge diet was able to supply 31 nutrients in precise amounts using only 33 ingredients". Prove that the manufacturer could have supplied the same 31 nutrients in precise amounts using only 31 of fewer ingredients.

12. The Birkhoff Polytope

(12.1) Definitions. Let σ be a permutation of the set $\{1, \dots, n\}$. The *permutation matrix* X^σ is an $n \times n$ matrix $X^\sigma = (\xi_{ij}^\sigma) : i, j = 1, \dots, n$, defined as follows:

$$\xi_{ij}^\sigma = \begin{cases} 1 & \text{if } \sigma(j) = i \\ 0 & \text{otherwise.} \end{cases}$$

An $n \times n$ matrix $X = \big(\xi_{ij}\big) : i, j = 1, \dots, n$ is called *doubly stochastic* provided it is non-negative $\xi_{ij} \geq 0$ for $i, j = 1, \dots, n$ and the sum of entries in every row and

every column is 1:

$$\sum_{i=1}^{n} \xi_{ij} = 1 \quad \text{for} \quad j = 1, \ldots, n \qquad \text{and} \qquad \sum_{j=1}^{n} \xi_{ij} = 1 \quad \text{for} \quad i = 1, \ldots, n.$$

The polyhedron B_n of all $n \times n$ doubly stochastic matrices is called the *Birkhoff Polytope.*

PROBLEMS.

1°. Prove that the set of integer doubly stochastic matrices is the set of permutation matrices.

2°. Prove that permutation matrices X^σ are extreme points of the Birkhoff Polytope.

The following remarkable result was established independently by G. Birkhoff in 1946 and by J. von Neumann in 1953.

(12.2) Birkhoff - von Neumann Theorem. *The vertices of the Birkhoff Polytope B_n are exactly the $n \times n$ permutation matrices.*

Proof. Because of Problem 2°, Section 12.1, it suffices to prove that if X is an extreme point of B_n, then $X = X^\sigma$ for some permutation σ. We prove this by induction on n. The case $n = 1$ is obvious. Suppose that $n > 1$. Let us consider the affine subspace $L \subset \mathbb{R}^{n^2}$ consisting of the $n \times n$ matrices $X = (\xi_{ij})$ such that

$$\sum_{i=1}^{n} \xi_{ij} = 1 \quad \text{for} \quad j = 1, \ldots, n \qquad \text{and} \qquad \sum_{j=1}^{n} \xi_{ij} = 1 \quad \text{for} \quad i = 1, \ldots, n.$$

We claim that $\dim L = (n-1)^2$. Indeed, if we arbitrarily fill the upper left $(n-1) \times (n-1)$ corner by numbers $\xi_{ij} : i, j = 1, \ldots, n-1$, there is a unique way to append it to a matrix from L: $\xi_{in} = 1 - \sum_{j=1}^{n-1} \xi_{ij}$ for $i = 1, \ldots, n-1$, $\xi_{nj} = 1 - \sum_{i=1}^{n-1} \xi_{ij}$ for $j = 1, \ldots, n-1$ and $\xi_{nn} = (2-n) + \sum_{i,j=1}^{n-1} \xi_{ij}$. In the space L, the polytope B_n is defined by n^2 linear inequalities $\xi_{ij} \geq 0$. If X is an extreme point of B_n, by Theorem 11.2, some $(n-1)^2$ of these inequalities must be active on X. In other words, $\xi_{ij} = 0$ for some $(n-1)^2$ entries of X. Clearly, there can not be a row containing zeros alone, and if every row contained at least two non-zero entries, the total number of zero entries would have been at most $n(n-2) < (n-1)^2$. Therefore, there must be a row, say, i_0 with $\xi_{i_0 j} = 0$ for all but one $j = j_0$. Now it is clear that $\xi_{i_0 j_0} = 1$ and that all other entries in the i_0-th row and in the j_0-th column must be zero. Crossing out the i_0-th row and the j_0-th column we get an $(n-1) \times (n-1)$ doubly stochastic matrix, which must be an extreme point of B_{n-1}, so we may apply the induction hypothesis. □

PROBLEMS.

1. Prove that $\dim B_n = (n-1)^2$, so B_n has an interior point in the subspace L, constructed in the proof.

2. Find the radius of the ball in L, centered at $\xi_{ij} = 1/n$ and touching the boundary of the polytope B_n.

3. Prove that the set $F = \{x \in B_n : \xi_{11} = 0\}$ is a face of B_n of dimension $(n-1)^2 - 1$ and that $G = \{x \in B_n : \xi_{11} = 1\}$ is a face of B_n of dimension $(n-2)^2$.

4°. Draw a picture of B_2.

5°. Let $U = (\zeta_{ij})$ be an $n \times n$ real orthogonal (that is, $UU^t = I$) or complex unitary (that is, $U\overline{U^t} = I$) matrix. Let $\beta_{ij} = |\zeta_{ij}|^2$. Prove that $B = (\beta_{ij})$ is a doubly stochastic matrix.

The problem of optimizing a linear function on the polytope B_n has an interesting combinatorial interpretation.

(12.3) The Assignment Problem. The *Assignment Problem* is formulated as follows: given an $n \times n$ matrix $C = (\gamma_{ij})$, find a permutation σ of the set $\{1, \ldots, n\}$ such that $\sum_{i=1}^{n} \gamma_{i\sigma(i)}$ is the maximum (or minimum) possible. A typical interpretation of this problem is as follows: There are n candidates to fill n positions. Let γ_{ij} be the "benefit" (or the "damage") brought by the assignment of the i-th candidate to the j-th position. We are seeking to maximize (or minimize) the total benefit (or damage). Because of Theorem 12.2, the Assignment Problem can be posed as a problem of finding the maximum (minimum) value of a linear function on the polytope B_n, in short, as a *linear programming problem*:

$$\text{Find } \max \sum_{ij=1}^{n} \gamma_{ij}\xi_{ij}$$

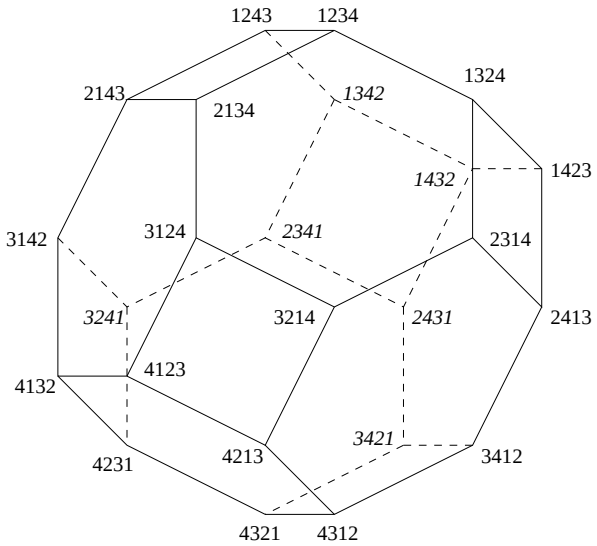
$$\text{Subject to: } \sum_{i=1}^{n} \xi_{ij} = 1 \text{ for } j = 1, \ldots, n; \qquad \sum_{j=1}^{n} \xi_{ij} = 1 \text{ for } i = 1, \ldots, n;$$

$$\xi_{ij} \geq 0 \text{ for } i, j = 1, \ldots, n.$$

13. The Permutation Polytope and the Schur-Horn Theorem

(13.1) Definition. Let us fix a point $x = (\xi_1, \ldots, \xi_n)$ in $\mathbb{R}^n$. For a permutation σ of the set $\{1, \ldots, n\}$, let $\sigma(x)$ be the vector $y = (\eta_1, \ldots, \eta_n)$, where $\eta_i = \xi_{\sigma^{-1}(i)}$. Let S_n be the symmetric group of all permutations of the set $\{1, \ldots, n\}$. Let us define the *permutation polytope* $P(x)$ by

$$P(x) = \operatorname{conv}\{\sigma(x) : \sigma \in S_n\}.$$



Permutation polytope $P(x)$ for $x=(1,2,3,4)$

Figure 16

PROBLEMS.

1° Prove that $\sigma(x) = X^\sigma x$, where X^σ is the permutation matrix corresponding to σ and that $(\sigma\tau)(x) = \sigma\big(\tau(x)\big)$ for every two permutations σ and τ.

2°. Let us interpret $\mathbb{R}^{n^2}$ as the space of $n \times n$ matrices X. Let us fix a vector $a \in \mathbb{R}^n$. Consider the linear transformation $T : \mathbb{R}^{n^2} \longrightarrow \mathbb{R}^n$ defined by $T(X) = Xa$. Prove that $T(B_n) = P(a)$, where $B_n \subset \mathbb{R}^{n^2}$ is the Birkhoff polytope.

3°. Prove that the permutation polytope $P(a)$, $a = (\alpha_1, \ldots, \alpha_n)$ lies in the affine hyperplane $x = (\xi_1, \ldots, \xi_n) : \xi_1 + \ldots + \xi_n = \alpha_1 + \ldots + \alpha_n$.

4. Suppose that not all the coordinates of a are equal. Prove that $\dim P(a) = n - 1$.

5°. Prove that $P(a)$ has $n!$ vertices if and only if the coordinates $\alpha_1, \ldots, \alpha_n$ are distinct.

6. Draw a picture of the permutation polytope $P(x)$ for $x = (1, 2, 3)$.

(13.2) Schur-Horn Theorem. *Let us fix a positive integer n and n real numbers $\lambda_1, \ldots, \lambda_n$. Let $l = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ be a vector.*

(1) *Let $A = (\alpha_{ij})$ be an $n \times n$ real symmetric (or complex Hermitian) matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the diagonal $a = (\alpha_{11}, \ldots, \alpha_{nn})$ lies in the permutation polytope $P(l)$: $a \in P(l)$ (Schur's Theorem).*

(2) *Let $a \in P(l)$ be a point from the permutation polytope. Then there exists an $n \times n$ real symmetric matrix $A = (\alpha_{ij})$ with the eigenvalues $\lambda_1, \ldots, \lambda_n$ and the diagonal $a = (\alpha_{11}, \ldots, \alpha_{nn})$ (Horn's Theorem).*

We will prove Schur's Theorem only (Part 1).

Proof of Part 1. Let $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ be the diagonal matrix with $\lambda_1, \dots, \lambda_n$ on the diagonal. Suppose that $A = (\alpha_{ij})$ is a real symmetric $n \times n$ matrix with the eigenvalues $\lambda_1, \dots, \lambda_n$ (the proof for complex Hermitian matrices is completely analogous). Then $A = UDU^t$ for some orthogonal matrix $U = (\zeta_{ij})$. Hence the diagonal entries of A can be written as

$$\alpha_{kk} = \sum_{i=1}^{n} \zeta_{ki}^2 \lambda_i.$$

Let $B = (\beta_{ij})$ be the $n \times n$ matrix such that $\beta_{ij} = \zeta_{ij}^2$. Hence we may write $a = Bl$, where a and l are interpreted as n-columns of real numbers. Since U is an orthogonal matrix, the matrix B is doubly stochastic (cf. Problem 5°, Section 12.2), that is, B is a non-negative matrix with all row and column sums equal to 1. By the Birkhoff - von Neumann Theorem (Theorem 12.2), B can be written as a convex combination of permutation matrices X^σ, $\sigma \in S_n$. Therefore we conclude that a is a convex combination of $\sigma(l) = X^\sigma l$, that is $a \in P(l)$ by Problem 1 of Section 13.1. □

For the proof of Part 2, see, A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

14. The Transportation Polyhedron

Let $G = (V, E)$ be a directed (finite) graph with the set of vertices $V = \{1, \dots, n\}$ and a set of edges E. Any two vertices $i, j \in V$ can be either connected by an edge $i \to j$, or two edges $i \to j$ and $j \to i$ going in the opposite directions or not connected at all. We assume that the graph has no loops $i \to i$.

Suppose further that to each vertex i, a real number α_i is assigned, which can be positive ("demand"), or negative ("supply"), or zero ("transit"). Suppose that to every edge $i \to j$ a number ξ_{ij} is assigned so that the following conditions are satisfied:

The balance requirement:

For every vertex $i \in V$

$$\sum_{j:\ (j \to i) \in E} \xi_{ji} - \sum_{j:\ (i \to j) \in E} \xi_{ij} = \alpha_i$$

No wrong way shipment:

For every edge $(i \to j) \in E$,

$$\xi_{ij} \geq 0.$$

An assignment of numbers ξ_{ij} satisfying the above requirements is called *a feasible flow* in G.

(14.1) Definition. Let us fix a graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges and a vector $a = (\alpha_1, \ldots, \alpha_n)$. Let us think of a feasible flow (ξ_{ij}) in G as a point in $\mathbb{R}^m$. The set of all feasible flows $x \in \mathbb{R}^m$ is called the *transportation polyhedron* and denoted $T(G, a)$.

PROBLEMS.

1. Prove that if the transportation polyhedron $T(G, a)$ is non-empty, then $\sum_{i=1}^n \alpha_i = 0$.

2°. Construct an example where $\sum_{i=1}^n \alpha_i = 0$ but the transportation polyhedron $T(G, a)$ is empty.

3°. Construct an example of an unbounded transportation polyhedron $T(G, a)$.

4. Suppose that G does not contain any directed cycle $i_1 \to i_2 \to \ldots \to i_l \to i_1$ for $l \geq 2$. Prove that $T(G, a)$ is a bounded polyhedron (polytope). It is called the *transportation polytope*.

(14.2) Proposition. *Let $x = \big(\xi_{ij} : (i \to j) \in E\big)$ be an extreme point of the transportation polyhedron $T(G, a)$. Let $S \subset E$ be the set of all edges $i \to j$ where $\xi_{ij} > 0$. Then S does not contain any cycle $v_1 - v_2 - \ldots - v_l - v_1 : l \geq 2$, where the vertices $v_k, v_{k+1} : k = 1, \ldots, l-1$ and v_l, v_1 are connected by an edge (of whatever orientation: $v_k \to v_{k+1}$ or $v_{k+1} \to v_k$).*

Proof. Suppose that there is such a cycle C consisting of edges with a strictly positive flow. Let us choose an $\epsilon > 0$ and let us construct two flows $y = (\eta_{ij})$ and $z = (\zeta_{ij})$ as follows:

$$\eta_{ij} = \begin{cases} \xi_{ij} & \text{if } (i \to j) \notin C \\ \xi_{ij} + \epsilon & \text{if } i = v_k, j = v_{k+1} \text{ or } i = v_l, j = v_1 \\ \xi_{ij} - \epsilon & \text{if } i = v_{k+1}, j = v_k \text{ or } i = v_1, j = v_l \end{cases}$$

and

$$\zeta_{ij} = \begin{cases} \xi_{ij} & \text{if } (i \to j) \notin C \\ \xi_{ij} - \epsilon & \text{if } i = v_k, j = v_{k+1} \text{ or } i = v_l, j = v_1 \\ \xi_{ij} + \epsilon & \text{if } i = v_{k+1}, j = v_k \text{ or } i = v_1, j = v_l. \end{cases}$$

In other words, we choose a direction of the cycle C (say, clockwise). To construct y, we increase the flow on the edges of the cycle that go clockwise by ϵ and decrease the flow on the edges of the cycle that go counterclockwise by ϵ. To construct z, we decrease the flow on the clockwise edges by ϵ and increase the flow on the

counterclockwise edges by ϵ, see Figure 17.

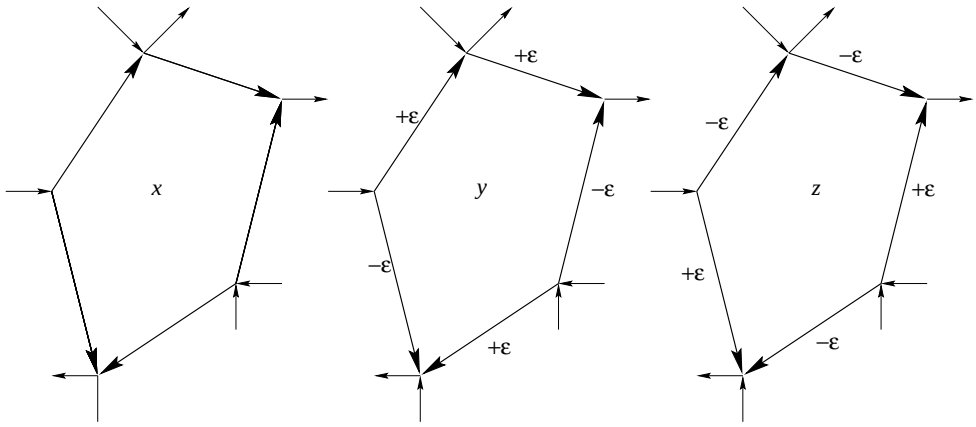


Figure 17

Then the flows y and z satisfy the balance condition. Furthermore, if $\epsilon > 0$ is small enough, y and z are non-negative flows and hence feasible flows. Finally, $x = (y + z)/2$, which proves that x can not be an extreme point of $T(G, a)$. □

(14.3) Definition. A graph without cycles is called a *forest*. A connected graph without cycles is called a *tree*.

PROBLEMS.

1°. Prove that a forest is a union of non-intersecting trees.

2. Prove that every (finite) forest has a vertex which is connected with at most one vertex of the forest.

Hint: Starting with a vertex of the forest, follow some path while there is still a way to go, without turning back. Since there are no cycles and the number of the vertices is finite, the path will end somewhere. The last vertex of the path is the vertex with the required property.

(14.4) Corollary. *Suppose that all demands/supplies α_i, $i = 1, \ldots, n$ are integer numbers. Then every extreme point of the transportation polyhedron $T(G, a)$ is an integer flow ξ_{ij}.*

Proof. Let $x = (\xi_{ij})$ be an extreme point. Proposition 14.2 asserts that the set $S \subset E$ with non-zero flows x_{ij} is a forest. We claim that once we know S, we can compute the flows ξ_{ij} from $\alpha_1, \ldots, \alpha_n$ by using addition and subtraction only. This, of course, would imply that all ξ_{ij} are integers.

If S is a forest with at least one edge, by Problem 2, Section 14.3, there is a vertex i such that there is only one edge with a non-zero flow incident to i. Let j be the other end of that edge. Clearly, if this edge is $i \to j$, we must have $\alpha_i < 0$ and $\xi_{ij} = -\alpha_i$. If this edge is $j \to i$, we must have $\alpha_i > 0$ and $\xi_{ji} = \alpha_i > 0$. Now we delete the vertex i with all edges of G incident to it, modify forest S accordingly,

and adjust the demand/supply vector: if we had $\xi_{ij} > 0$, we let new $\alpha_j := \alpha_j + \alpha_i$, if we had $\xi_{ji} > 0$ we let new $\alpha_j = \alpha_j - \alpha_i$. Hence we get a new graph G' with $n-1$ vertices and integer demands/supplies, a new forest S' of G' and a new feasible flow ξ'_{ij} in G' such that S' is the set of edges where ξ'_{ij} are strictly positive. We proceed as above, until there are no edges in the forest. At that moment the flow ξ_{ij} is determined completely. □

PROBLEMS.

1. Deduce the Birkhoff - von Neumann Theorem (Theorem 12.2) from Corollary 14.4 as follows: Consider the graph G with $2n$ vertices $1, \dots, n$ and $1', \dots, n'$. Let $\alpha_i = 1$ be the demand for $i = 1, \dots, n$ and let $a_{i'} = -1$ be the supply for $i' = 1', \dots, n'$. Let $i' \to j$, $i' = 1', \dots, n'$, $j = 1, \dots, n$ be the edges of G. Prove that the feasible flows $x_{i'j}$ are the doubly stochastic matrices and that the integral feasible flows are the permutation matrices.

2. Let us fix positive integers m and n and let us interpret a real $m \times n$ matrix as a point in $\mathbb{R}^{mn}$. Let $a = (\alpha_1, \dots, \alpha_m)$ and $b = (\beta_1, \dots, \beta_n)$ be two vectors of positive integers and let $P(a, b)$ be the set of all non-negative $m \times n$ matrices with the row sums $\alpha_1, \dots, \alpha_m$ and the column sums $\beta_1, \dots, \beta_n$. Prove that $P(a, b)$ is a bounded polyhedron (polytope) and that every vertex of $P(a, b)$ is an integer matrix.

(14.5) The transportation problem.

Suppose that γ_{ij} are (usually non-negative) costs on the edges of the graph $G = (V, E)$. The problem of finding a feasible flow $x \in T(G, a)$ minimizing the total cost $\sum_{i \to j \in E} \gamma_{ij} \xi_{ij}$ is called the *Transportation Problem.*

From Proposition 14.2 and Theorem 10.2 (Part 1) we deduce that if the optimal flow is unique, the set of edges $i \to j$ where the flow is positive must form a forest in G. Furthermore, by Corollary 10.3 we conclude that if $T(G, a)$ is bounded and non-empty, then there is an optimal flow with that property.

PROBLEM.

1. Suppose that the transportation polyhedron $T(G, a)$ is non-empty and that the cost γ_{ij} is strictly positive for every edge $i \to j$ in E. Prove that there exists an optimal flow in the transportation problem and that the set of all optimal flows is a compact polyhedron (polytope), which is a face of the transportation polyhedron $T(G, a)$. Deduce that there exists an optimal solution such that the set of all edges where the flow is positive forms a forest in G.

15. Convex Cones

We introduce convex cones, an interesting and important type of convex sets.

(15.1) Cones, conic hulls and extreme rays. Let V be a vector space. A set $K \subset V$ is called a *cone* if $0 \in K$ and $\lambda x \in K$ for every $\lambda \geq 0$ and every $x \in K$. The cones we will be dealing with are convex. Alternatively, we can say that $K \subset V$ is

a convex cone if $0 \in K$ and if for any two points $x, y \in K$ and any two numbers $\alpha, \beta \geq 0$, the point $z = \alpha x + \beta y$ is also in K.

Given points $x_1, \ldots, x_m \in V$ and non-negative numbers $\alpha_1, \ldots, \alpha_m$, the point $x = \sum_{i=1}^{m} \alpha_i x_i$ is called a *conic* combination of the points $x_1, \ldots, x_m$. The set $\mathrm{co}(S)$ of all conic combinations of points from a set $S \subset V$ is called the *conic hull* of the set S. The conic hull $\mathrm{co}(x)$ of a non-zero point $x \in V$ is called the *ray* spanned by x.

Let $K \subset V$ be a cone and let $K_1 \subset K$ be a ray. We say that K_1 is an *extreme ray* of K if for any $u \in K_1$ and any $x, y \in K$, whenever $u = (x + y)/2$, we must have $x, y \in K_1$.

Let K be a cone and let $x \in K$ be a point. If $K_1 = \mathrm{co}(x)$ is an extreme ray of K, we say that x *spans* an extreme ray (of K).

PROBLEMS.

1°. Prove that $\mathrm{co}(S)$ is the smallest convex cone containing the set $S \subset V$, that is the intersection of all convex cones in V that contain S.

2°. Let $K_1, K_2 \subset V$ be convex cones. Prove that the intersection $K_1 \cap K_2$ and the Minkowski sum $K_1 + K_2 = \{x + y : x \in K_1, y \in K_2\}$ are convex cones.

3. Construct an example of two convex closed cones $K_1, K_2 \subset \mathbb{R}^3$ such that $K_1 + K_2$ is not closed.

4. Prove that the closure of a convex cone in $\mathbb{R}^d$ is a convex cone.

5. Prove that 0 is an extreme point of a convex cone K if K does not contain a straight line and construct an example of a convex cone $K \subset \mathbb{R}^2$, such that K contains a straight line but 0 is an extreme point of K.

6. Let $S \subset \mathbb{R}^d$ be a set. Prove that every point $x \in \mathrm{co}(S)$ is a conic combination of some d points from S.

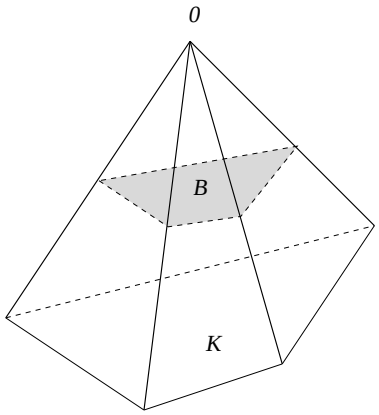
We need a few technical lemmas, adjusting our results to convex cones.

(15.2) Lemma. *Let $K \subset V$ be a cone and let $H \subset V$ be an affine hyperplane isolating K and such that $K \cap H \neq \emptyset$. Then $0 \in H$.*

Proof. Assume that for some (non-zero) linear function $f : V \longrightarrow \mathbb{R}$ and some number α we have $H = \{x : f(x) = \alpha\}$ and that $K \subset H_-$. Since $0 \in K$, we get $\alpha \geq 0$. Suppose that $\alpha > 0$ and let $x \in K$ be a point such that $f(x) = \alpha$. Then, for $\lambda > 1$ we have $f(\lambda x) = \lambda\alpha > \alpha$ and $\lambda x \in K$, which contradicts the assumption that $K \subset H_-$. Hence we must have $\alpha = 0$ and $0 \in H$. □

PROBLEM. 1°. Prove that a non-empty face of a cone is a cone.

(15.3) Definition. Let $K \subset V$ be a cone. A set $B \subset K$ is called a *base* of K if $0 \notin B$ and for every point $u \in K, u \neq 0$ there is a unique representation $u = \lambda v$ with $v \in B$ and $\lambda > 0$.



A base B of the cone K

Figure 18

(15.4) Lemma. *Let K be a cone with a convex base B and let $u \in K$, $u \neq 0$ be a non-zero point from K. Then u spans an extreme ray of K if and only if $u = \lambda v$, where $\lambda > 0$ and v is an extreme point of B.*

Proof. Suppose that u spans an extreme ray of K. Let $v \in B$ be a point such that $u = \lambda v$ for $\lambda > 0$. Suppose that $v = (v_1 + v_2)/2$. Then $u = (u_1 + u_2)/2$, where $u_1 = \lambda v_1$ and $u_2 = \lambda v_2$. Since u spans an extreme ray, we must have $u_1 = \mu_1 u$ and $u_2 = \mu_2 u$ for some $\mu_1, \mu_2 \geq 0$. Then $u_1 = (\mu_1 \lambda)v$ and $u_2 = (\mu_2 \lambda)v$. Since B is a base, we must have $v_1 = v_2 = v$, so v is an extreme point of B.

Suppose that $u = \lambda v$, where $\lambda > 0$ and v is an extreme point of B. Let us show that u spans an extreme ray of K. Suppose that $u = (u_1 + u_2)/2$. Then $u_1 = \lambda_1 v_1$ and $u_2 = \lambda_2 v_2$ for some $v_1, v_2 \in B$ and some non-negative λ_1, λ_2. Without loss of generality, we can assume that $\lambda_1, \lambda_2 > 0$. Then we can write $u = (\lambda_1 v_1 + \lambda_2 v_2)/2 = \beta(\alpha_1 v_1 + \alpha_2 v_2)$, where $\beta = (\lambda_1 + \lambda_2)/2$, $\alpha_1 = \lambda_1/(\lambda_1 + \lambda_2)$, and $\alpha_2 = \lambda_2/(\lambda_1 + \lambda_2)$. Since B is a base, we must have $\alpha_1 v_1 + \alpha_2 v_2 = v$. We note, that $\alpha_1, \alpha_2 > 0$ and that $\alpha_1 + \alpha_2 = 1$. If $v_1 \neq v_2$, it follows that v lies inside the interval $[v_1, v_2]$, which contradicts the assumption that v is an extreme point of B. Therefore, we must have $v_1 = v_2$, so u spans an extreme ray of K. □

We obtain a conic version of the Krein-Milman Theorem.

(15.5) Corollary. *Let $K \subset \mathbb{R}^d$ be a cone with a compact base. Then every point $u \in K$ can be written as a conic combination*

$$u = \sum_{i=1}^{m} \lambda_i u_i, \quad \lambda_i \geq 0 : \ i = 1, \ldots, m,$$

where points u_i span extreme rays of K.

Proof. Let B be a base of K. Let us write $u = \lambda v$, where $v \in B$. By the Krein-Milman Theorem (Theorem 10.3), we can express v as a convex combination of extreme points $v_1, \ldots, v_m$ of B. Then u is a conic combination of

$u_1 = \lambda_1 v_1, \ldots, u_m = \lambda_m v_m$. By Lemma 15.4, the points u_i span extreme rays of K. $\square$

(15.6) Lemma. *Let $K \subset \mathbb{R}^d$ be a cone with a compact base. Then K is closed.*

Proof. Let $B \subset \mathbb{R}^d$ be a compact base of K. Consider a sequence of points $\{x_n \in K\}$ such that $\lim x_n = x \in \mathbb{R}^d$. Our goal is to prove that $x \in K$. We can write $x_n = \lambda_n v_n$, where $v_n \in B$ and $\lambda_n \geq 0$. Since $0 \notin B$ and B is compact we have $\|v_n\| \geq r$ for some $r > 0$ and all n. Similarly, since $\{x_n\}$ is a converging sequence, $\|x_n\| \leq R$ for some R and all n. Therefore, $\lambda_n \leq R/r$. Since B is compact and the sequence λ_n is bounded, we can choose a subsequence $\{n_k\}$ such that $\lim \lambda_{n_k} = \lambda$ and $\lim v_{n_k} = v \in B$. Then $x = \lim x_{n_k} = \lambda v \in K$ and the proof follows. $\square$

PROBLEMS.

1. Let $K \subset \mathbb{R}^d$ be a cone with a compact base. Prove that 0 is a face of K.
2. Let $K \subset \mathbb{R}^n$ be a cone with a compact base B and let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation such that $(\ker T) \cap B = \emptyset$. Prove that $T(K)$ is a cone with a compact base in $\mathbb{R}^m$.
3. Construct an example of a compact set $A \subset \mathbb{R}^2$ such that $\mathrm{co}(A)$ is not closed.

16. The Moment Curve and the Moment Cone

(16.1) The moment curve. Let us consider the space $\mathbb{R}^{d+1}$ with the coordinates $x = (\xi_0, \xi_1, \ldots, \xi_d)$ (we start with the zeroth coordinate). Given real numbers $\alpha < \beta$, the curve

$$g(\tau) = (1, \tau, \tau^2, \ldots, \tau^d) \in \mathbb{R}^{d+1} \quad \text{for} \quad \alpha \leq \tau \leq \beta$$

is called *the moment curve*. Hence $g(\tau)$ lies in the affine hyperplane $\xi_0 = 1$ in $\mathbb{R}^{d+1}$.

Let $f(x) = \langle c, x \rangle$ be a linear function, where $c = (\gamma_0, \gamma_1, \ldots, \gamma_d)$. Then

$$f\big(g(\tau)\big) = \gamma_d \tau^d + \gamma_{d-1}\tau^{d-1} + \ldots + \gamma_1 \tau + \gamma_0$$

is a polynomial of degree d.

PROBLEMS.

1. Prove that each linear (that is, passing through 0) hyperplane $H \subset \mathbb{R}^{d+1}$ intersects the moment curve $g(\tau)$ in at most d points.
2. Let $\delta_1 < \delta_2 < \ldots < \delta_d$ be real numbers. Consider the curve in $\mathbb{R}^d$

$$h(\tau) = \big(\exp\{\delta_1 \tau\}, \exp\{\delta_2 \tau\}, \ldots, \exp\{\delta_d \tau\}\big) \quad \text{for } \alpha \leq \tau \leq \beta.$$

Prove that each affine hyperplane $H \subset \mathbb{R}^d$ intersects $h(\tau)$ in at most d points.

3. Let S^1 be the circle in the natural parameterization $S^1 = \{0 \leq \tau \leq 2\pi\}$. Suppose that $d = 2k$ is even and let $h : S^1 \longrightarrow \mathbb{R}^d$ be the closed curve

$$h(\tau) = \big(\sin\tau, \cos\tau, \sin 2\tau, \cos 2\tau, \ldots, \sin k\tau, \cos k\tau\big), \quad 0 \leq \tau \leq 2\pi.$$

Prove that each affine hyperplane $H \subset \mathbb{R}^d$ intersects the curve $h(\tau)$ in at most d points.

4. Suppose that d is odd. Prove that one can not embed the circle S^1 into $\mathbb{R}^d$ in such a way that every affine hyperplane intersects the circle in not more than d points.

5. Consider the set "Y" in the plane (three intervals having one common point). Prove that for any d one can not embed Y into $\mathbb{R}^d$ in such a way that every affine hyperplane intersects Y in not more than d points.

(16.2) Definition. The *moment cone*

$$M_{d+1} = \operatorname{co}\big(g(\tau) : \alpha \le \tau \le \beta\big) \subset \mathbb{R}^{d+1}$$

is the conic hull of the curve $g(\tau)$. Sometimes we write $M_{d+1}[\alpha, \beta]$ instead of M_{d+1}.

PROBLEMS.

1°. Prove that $\operatorname{conv}\big(g(\tau) : \alpha \le \tau \le \beta\big)$ is a convex compact base of $M_{d+1}[\alpha, \beta]$

2°. Prove that $\dim M_{d+1} = d + 1$.

(16.3) Lemma. *The moment cone M_{d+1} is closed.*

Proof. Follows from Lemma 15.6 and Problem 1 of Section 16.2. □

(16.4) Proposition. *Let $u \in \partial M$ be a point on the boundary of M_{d+1}. Let us write u as a conic combination of points on the curve $g(\tau)$:*

$$u = \sum_{j=1}^{m} \lambda_j g(\tau_j), \quad \text{where } \lambda_j > 0 \text{ for } j = 1, \dots, m \quad \text{and} \quad \alpha \le \tau_1 < \dots < \tau_m \le \beta.$$

Then $m \le (d+2)/2$. Furthermore, if $m = (d+2)/2$, then d is even and $\tau_1 = \alpha$, $\tau_m = \beta$.

Proof. By Problem 2 of Section 16.2, $\operatorname{int} M_{d+1} \ne \emptyset$, so by Theorem 9.7 there exists an affine hyperplane H that contains u and isolates M_{d+1}. Hence by Lemma 15.2, H contains the origin, so $H = \big\{x : \langle c, x\rangle = 0\big\}$ for some $c = (\gamma_0, \gamma_1, \dots, \gamma_d) \ne 0$. We have $\langle c, x\rangle \ge 0$ for $x \in M_{d+1}$ and $\langle c, u\rangle = 0$. In particular, $\langle c, g(\tau)\rangle \ge 0$ for all $\alpha \le \tau \le \beta$. Let

$$p(\tau) = \langle c, g(\tau)\rangle = \gamma_0 + \gamma_1\tau + \dots + \gamma_d\tau^d.$$

Hence $p(\tau)$ is a polynomial of degree d which is non-negative on the interval $[\alpha, \beta]$. Furthermore, since $\langle c, u\rangle = 0$ and $\lambda_j > 0$ we must have $p(\tau_j) = 0$ for $j = 1, \dots, m$. Suppose that $\tau^* \in (\alpha, \beta)$ is a root of p, which lies strictly inside the interval (α, β). Then the multiplicity of the root must be an even number, since otherwise $p(\tau)$ would change sign in a neighborhood of τ^*. The only roots of multiplicity one can be $\tau^* = \alpha$ and $\tau^* = \beta$. The total number of roots of p, counting multiplicities, is at most d. If all the roots of p are strictly inside (α, β) then $2m \le d$ and $m \le d/2$. If only one endpoint of the interval $[\alpha, \beta]$ is a root of p then there are $m - 1$ roots inside (α, β) and $2(m-1) + 1 \le d$, so $m \le (d+1)/2$. If both endpoints α and β are roots of p then there are $(m-2)$ roots inside (α, β), so $2 + 2(m-2) \le d$, so $m \le (d+2)/2$. Thus in any case, $m \le (d+2)/2$. □

(16.5) Corollary. *Let*

$$m = \begin{cases} (d+1)/2 & \text{if } d \text{ is odd} \\ (d+2)/2 & \text{if } d \text{ is even.} \end{cases}$$

Let $u \in M_{d+1}[\alpha, \beta]$ be a point. Then u can be represented as a conic combination of m (or fewer) points on the curve $g(\tau)$:

$$u = \sum_{i=1}^{m} \lambda_i g(\tau_i)$$

where $\lambda_i \geq 0$ for $i = 1, \ldots, m$ and $\alpha \leq \tau_1 < \ldots < \tau_m \leq \beta$.

Proof. If $u \in \partial M_{d+1}$, the result follows by Proposition 16.4. Suppose that $u \in \operatorname{int} M_{d+1}[\alpha, \beta]$. Let $t \geq 0$ be a parameter, and let us "shrink" the interval $[\alpha, \beta] \longrightarrow [\alpha + t, \beta - t]$ as t grows. Let us consider $M_{d+1}[\alpha + t, \beta - t]$. When $t = (\beta - \alpha)/2$, the curve g consists of a single point, so the cone $M_{d+1}[\alpha + t, \beta - t]$ consists of a single ray. If the point u is still in the cone, the proof follows. Otherwise, there is a value t^* such that $u \in \partial M_{d+1}[\alpha + t^*, \beta - t^*]$. Now we use Proposition 16.4. □

17. An Application: "Double Precision" Formulas for Numerical Integration

Corollary 16.5 has a somewhat unexpected application.

(17.1) Proposition. *Let us fix an interval $[\alpha, \beta]$ and a non-negative continuous density $\rho(\tau)$ on $[\alpha, \beta]$. Then, for any positive integer m, there exist m points $\tau_1^*, \ldots, \tau_m^*$ on the interval $[\alpha, \beta]$ and m non-negative numbers $\lambda_1, \ldots, \lambda_m$ such that*

$$\int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \sum_{i=1}^{m} \lambda_i f(\tau_i^*)$$

for any polynomial f of degree at most $2m - 1$.

Proof. Let $d = 2m - 1$. Let $u = (\xi_0, \ldots, \xi_d) \in \mathbb{R}^{d+1}$, where $\xi_i = \int_\alpha^\beta \tau^i \rho(\tau)\, d\tau$ for $i = 0, \ldots, d$. Let us prove that $u \in M_{d+1}[\alpha, \beta]$. Indeed, since ρ is continuous, u can be written as a limit of Riemann sums:

$$u = \lim_{N \longrightarrow +\infty} \frac{\beta - \alpha}{N} \sum_{i=1}^{N} \rho(\tau_i) g(\tau_i),$$

where $\tau_1, \ldots, \tau_N$ are equally spaced points on $[\alpha, \beta]$. Since every Riemann sum is in the cone $M_{d+1}[\alpha, \beta]$, and by Lemma 16.3 the moment cone is closed, we get that

$u \in M_{d+1}$. Therefore, by Corollary 16.5, we can write u as a conic combination of m points on $g(\tau)$:

$$u = \sum_{i=1}^{m} \lambda_j g(\tau_i^*),$$

where $\lambda_i \geq 0$ and $\alpha \leq \tau_1^* < \ldots < \tau_m^* \leq \beta$. Now, let $f(\tau) = \gamma_d \tau^d + \ldots + \gamma_0$ be a polynomial of degree at most d. Let $c = (\gamma_0, \ldots, \gamma_d) \in \mathbb{R}^{d+1}$. Then

$$f(\tau_i^*) = \langle c, g(\tau_i^*) \rangle \text{ for } \; i = 1, \ldots, m \quad \text{and} \int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \langle c, u \rangle.$$

The proof follows. □

Formulas for numerical integration

$$\int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \sum_{i=1}^{m} \lambda_i f(\tau_i^*)$$

that are exact for polynomials f of degree up to $d = 2m - 1$ are often called the *double precision* integration formulas (for obvious reasons). The proof of Proposition 17.1 explains the term "Moment cone". The points of $M_{d+1}[\alpha, \beta]$ correspond to the *moments* of non-negative densities $\rho(\tau)$ on $[\alpha, \beta]$:

$$x = \left(\int_\alpha^\beta \rho(\tau)\, d\tau, \int_\alpha^\beta \tau\rho(\tau)\, d\tau, \ldots, \int_\alpha^\beta \tau^d \rho(\tau)\, d\tau \right).$$

As ρ varies, x ranges over the points in $M_{d+1}[\alpha, \beta]$.

PROBLEMS.

1. Prove that there are no m points $\tau_1^*, \ldots, \tau_m^*$ on the interval $[0, 1]$ and m real numbers $\lambda_1, \ldots, \lambda_m$ such that

$$\int_0^1 f(\tau)\, d\tau = \sum_{i=1}^{m} \lambda_i f(\tau_i^*)$$

for any polynomial of degree $2m$.

2. Prove that for every interval $[\alpha, \beta]$, every strictly positive continuous density $\rho(\tau)$ on $[\alpha, \beta]$, and every positive integer m there is only one set of points $\tau_1^*, \ldots, \tau_m^*$ on the interval $[\alpha, \beta]$ and only one set of non-negative numbers $\lambda_1, \ldots, \lambda_m$ such that

$$\int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \sum_{i=1}^{m} \lambda_i f(\tau_i^*)$$

for every polynomial of degree at most $2m - 1$.

3^*. A function

$$f(\tau) = \gamma_0 + \sum_{k=1}^{d} (\alpha_k \sin k\tau + \beta_k \cos k\tau) \quad \text{for} \quad 0 \le \tau \le 2\pi$$

is called a *trigonometric polynomial* of degree at most d. Let ρ be a non-negative continuous function on $[0, 2\pi]$ such that $\rho(0) = \rho(2\pi)$. Prove that there exist $d+1$ points $0 \le \tau_0^* < \ldots < \tau_d^* < 2\pi$ and $d+1$ non-negative numbers $\lambda_0, \ldots, \lambda_d$, such that the formula

$$\int_0^{2\pi} f(\tau)\rho(\tau)\, d\tau = \sum_{i=0}^{d} \lambda_i f(\tau_i^*)$$

is exact for any trigonometric polynomial of degree at most d.

Hint: Use Problem 3 of Section 16.1.

4^*. Let us fix $d = 2m$ distinct real numbers $\delta_1, \ldots, \delta_d$. A function

$$f(\tau) = \sum_{i=1}^{d} \alpha_i \exp\{\delta_i \tau\}$$

is called an *exponential polynomial* with exponents $\delta_1, \ldots, \delta_d$. Let ρ be a non-negative continuous density on the interval $[\alpha, \beta]$. Prove that there exist m points $\tau_1^*, \ldots, \tau_m^*$ on the interval $[\alpha, \beta]$ and m non-negative numbers $\lambda_1, \ldots, \lambda_m$, such that

$$\int_\alpha^\beta f(\tau)\rho(\tau)\, d\tau = \sum_{i=1}^{m} \lambda_i f(\tau_i^*)$$

for any exponential polynomial with exponents $\delta_1, \ldots, \delta_d$.

Hint: Use Problem 2 of Section 16.1.

Let us consider some small examples of double precision formulas for integrating $\int_0^1 f(\tau)\, d\tau$.

(17.2) Example. The formula that uses one node and is exact for polynomials of degree at most 1 is

$$\int_0^1 f(\tau)\, d\tau = f\Big(\frac{1}{2}\Big).$$

PROBLEM.

1°. Prove that the formula of Example 17.2 is indeed exact on polynomials of degree at most 1 and that this is the only such a formula.

(17.3) Example. Let us find formulas that use two nodes and and are exact on polynomials of degree at most 2. In the xy plane, let us consider the parabola arc $\big\{(\tau) = (\tau, \tau^2) : 0 \le \tau \le 1\big\}$. Let $u = (1/2, 1/3)$. There are infinitely many formulas

$$\int_0^1 f(\tau)\, d\tau = \lambda_1 f(\tau_1^*) + \lambda_2 f(\tau_2^*)$$

that are exact on polynomials of degree 2. The necessary and sufficient condition for $0 \le \tau_1^* < \tau_2^* \le 1$ is that the interval $[g(\tau_1^*), g(\tau_2^*)]$ on the xy plane contains u. Then λ_1 and λ_2 are found from $u = \lambda_1 g(\tau_1^*) + \lambda_2 g(\tau_2^*)$. For example, if we choose $\tau_1^* = 0$ then $\tau_2^* = 2/3$ and we get a formula

$$\int_0^1 f(\tau)\, d\tau = \frac{1}{4} f(0) + \frac{3}{4} f\Big(\frac{2}{3}\Big).$$

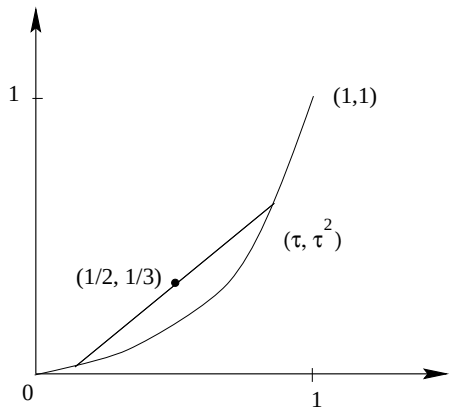


Figure 19

PROBLEM.

1. Find the one-parametric family of formulas with two nodes which are exact on polynomials of degree 2.

(17.4) Example. Here is a formula with two nodes which is exact for polynomials of degree at most 3:

$$\int_0^1 f(\tau)\, d\tau = \frac{1}{2} f\Big(\frac{1}{2} - \frac{1}{2\sqrt{3}}\Big) + \frac{1}{2} f\Big(\frac{1}{2} + \frac{1}{2\sqrt{3}}\Big).$$

PROBLEM.

1. Prove that the above formula is indeed exact on polynomials of degree at most 3 and that this is the only one such a formula which uses two nodes.

18. The Cone of Non-Negative Polynomials

(18.1) Definition. Let us interpret the space $\mathbb{R}^{d+1}$ as the space of all polynomials in one variable τ of degree at most d: a polynomial $\gamma_0+\gamma_1\tau+\ldots+\gamma_d\tau^d$ is represented by the point $(\gamma_0,\ldots,\gamma_d)$. Let us fix numbers $\alpha<\beta$ and let $K_+[\alpha,\beta]\subset\mathbb{R}^{d+1}$ be the set of all polynomials p that are non-negative on the interval $[\alpha,\beta]$: $p(\tau)\geq 0$ for all $\tau\in[\alpha,\beta]$. Sometimes we write K_+ instead of $K_+[\alpha,\beta]$.

PROBLEMS.

1°. Prove that $K_+[\alpha,\beta]$ is a convex closed cone in $\mathbb{R}^{d+1}$.

2°. Prove that $K_+[\alpha,\beta]$ has a non-empty interior.

3. Prove that $p\in\partial K_+$ if and only if $p(\tau)\geq 0$ for all $\tau\in[\alpha,\beta]$ and $p(\tau_0)=0$ for some $\tau_0\in[\alpha,\beta]$.

(18.2) Proposition. *The cone $K_+[\alpha,\beta]$ has a compact base. A polynomial $p\in K_+[\alpha,\beta]$ spans an extreme ray of $K_+[\alpha,\beta]$ if and only if the polynomial p has one of the following types:*

$$p(\tau)=\delta\prod_{i=1}^{k}(\tau-\tau_i)^2,\quad 2k=d$$

$$p(\tau)=\delta(\tau-\alpha)(\beta-\tau)\prod_{i=1}^{k}(\tau-\tau_i)^2,\quad 2k+2=d$$

for d even,

$$p(\tau)=\delta(\tau-\alpha)\prod_{i=1}^{k}(\tau-\tau_i)^2,\quad 2k+1=d$$

$$p(\tau)=\delta(\beta-\tau)\prod_{i=1}^{k}(\tau-\tau_i)^2,\quad 2k+1=d$$

for d odd, where $\delta>0$ and $\tau_1,\ldots,\tau_k$ are (not necessarily distinct) points from the interval $[\alpha,\beta]$.

Proof. Let

$$B=\Big\{q\in K_+:\int_\alpha^\beta q(\tau)\,d\tau=1\Big\}.$$

Obviously, B is non-empty convex closed set. It is also clear that for any $p\in K_+$, $p\neq 0$ there is a unique representation $p=\lambda q$ for $q\in B$ and $\lambda>0$. Take

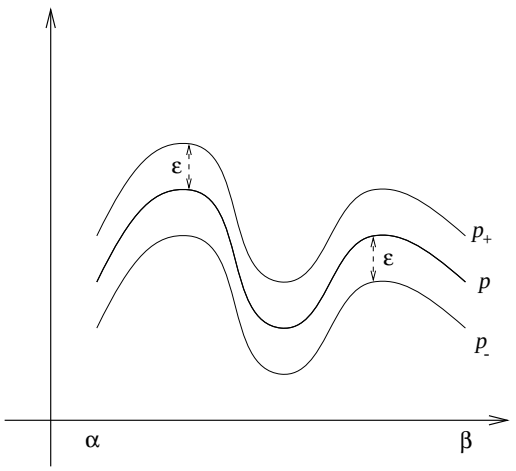
$$\lambda=\int_\alpha^\beta p(\tau)\,d\tau.$$

It remains to show that B is compact. Let us define a norm $\|\cdot\|_1$ on $\mathbb{R}^{d+1}$:

$$\|p\|_1 = \int_\alpha^\beta |p(\tau)|\, d\tau.$$

Let $\epsilon > 0$ be the minimum value of the continuous function $\|p\|_1$ on the unit sphere $S = \{p : \|p\| = 1\}$, where $\|\cdot\|$ is the usual Euclidean norm. Then $\|p\| \le 1/\epsilon$ for all $p \in B$, so B is compact. Hence B is a compact base of K_+.

Let $p \in K_+$ and suppose that $\mathrm{co}(p)$ is an extreme ray of K_+. If the $\deg p \le d$, we can write $p = (p_+ + p_-)/2$, where $p_+ = p + \epsilon(\tau - \alpha)p$ and $p_- = p - \epsilon(\tau - \alpha)p$. If $\epsilon > 0$ is sufficiently small than $p_+, p_- \in K_+[\alpha, \beta]$ and p_- and p_+ are not proportional to p, which is a contradiction. Hence $\deg p = d$. We can factor $p = qr$, where $q, r \in K_+$ and r is a polynomial without any root on the interval $[\alpha, \beta]$. In particular, r is separated from 0 on the interval. Let us choose a sufficiently small $\epsilon > 0$ such that $r_- = r - \epsilon$ and $r_+ = r + \epsilon$ are both non-negative on $[\alpha, \beta]$. We can write $p = (p_+ + p_-)/2$, where $p_+ = qr_+$ and $p_- = qr_-$. Since $\mathrm{co}(p)$ is an extreme ray, r_+ and r_- must be proportional to r and, therefore, r must be a constant. Summarizing, p has d roots on the interval $[\alpha, \beta]$. Finally, the multiplicity of every root of p which lies inside (α, β) must be even, since otherwise p changes its sign in a neighborhood of the root. So the polynomials that span extreme rays of K_+ must have the required structure.



Decomposition $p=(p_+ + p_-)/2$ of a positive polynomial

Figure 20

It remains to prove that the polynomials having the required type indeed span extreme rays of K_+. Suppose that $p = (p_1 + p_2)/2$ for some $p_1, p_2 \in K_+$. We claim that every root τ^* of p of multiplicity m must be a root of both p_1 and p_2 of multiplicity at least m. Otherwise, for one of the polynomials p_i, $i = 1, 2$ we'll have $p_i(\tau^*) = \ldots = p_i^{(k-1)}(\tau^*) = 0$, $p_i^{(k)}(\tau^*) < 0$ and for the other polynomial we'll have $p_i(\tau^*) = \ldots = p_i^{(k-1)}(\tau^*) = 0$, $p_i^{(k)}(\tau^*) > 0$ for some $0 \le k \le m$.

Therefore, one of the polynomials p_1 or p_2 would turn negative for some $\tau \in [\alpha, \beta]$ in a small neighborhood of the root τ^*. Since p has d roots, both p_1 and p_2 must be proportional to p.

□

(18.3) Corollary. *Let $p(\tau)$ be a polynomial such that $p(\tau) \geq 0$ for all $\tau \in [\alpha, \beta]$.*

If $d = 2k$ is even then

$$p(\tau) = (\tau - \alpha)(\beta - \tau) \sum_{i \in I} q_i^2(\tau) + \sum_{j \in J} q_j^2(\tau)$$

for some polynomials q_i, q_j. Furthermore, we can choose q_i, q_j is such a way that $\deg q_i = k - 1$ for $i \in I$, $\deg q_j = k$ for $j \in J$ and all roots of q_i, q_j are real and belong to the interval $[\alpha, \beta]$.

If $d = 2k + 1$ is odd then

$$p(\tau) = (\tau - \alpha) \sum_{i \in I} q_i^2(\tau) + (\beta - \tau) \sum_{j \in J} q_j^2(\tau)$$

for some polynomials q_i, q_j. Furthermore, we can choose q_i, q_j in such a way that $\deg q_i = \deg q_j = k$ for $i \in I$, $j \in J$ and all roots of q_i, q_j are real and belong to the interval $[\alpha, \beta]$.

Proof. Follows from Corollary 15.5 and Proposition 18.2. □

PROBLEMS.

1*. Prove that a polynomial $p(\tau)$ of degree d, which is non-negative on $[\alpha, \beta]$, has a unique representation

$$p(\tau) = \begin{cases} \delta \prod_{i=1}^{k} (\tau - \tau_{2i-1})^2 + \gamma(\tau - \alpha)(\beta - \tau) \prod_{i=1}^{k-1} (\tau - \tau_{2i})^2 & \text{if } \ d = 2k \\ \delta(\tau - \alpha) \prod_{i=1}^{k} (\tau - \tau_{2i})^2 + \gamma(\beta - \tau) \prod_{i=1}^{k} (\tau - \tau_{2i-1})^2 & \text{if } \ d = 2k + 1, \end{cases}$$

where $\gamma, \delta > 0$ and $\alpha \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_{d-1} \leq \beta$,
see S. Karlin and W.J. Studden, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Wiley, Interscience, 1966.

2. Let $K_+[0, +\infty) \subset \mathbb{R}^{d+1}$ be the set of all polynomials $p(\tau)$ of degree at most d such that $p(\tau) \geq 0$ for all $\tau \geq 0$. Prove that $K_+[0, +\infty)$ is a convex closed cone with a compact base and that the polynomials that span extreme rays of $K_+[0, +\infty)$ are

$$p(\tau) = \delta \prod_{i=1}^{k} (\tau - \tau_i)^2 \ \ (2k \leq d) \quad \text{and} \quad p(\tau) = \delta\tau \prod_{i=1}^{k} (\tau - \tau_i)^2 \ \ (2k + 1 \leq d),$$

where $\delta > 0$ and $\tau_i \geq 0$ for $i = 1, \ldots, k$.

Deduce that every polynomial p which is non-negative on $[0, +\infty)$ can be represented in the form

$$p(\tau) = \tau \sum_{i \in I}^{m} q_i^2(\tau) + \sum_{j \in J} q_j^2(\tau),$$

where q_i and q_j are polynomials with all roots real and non-negative.

3. Let $K_+(-\infty, +\infty) \subset \mathbb{R}^{d+1}$ be the set of all polynomials $p(\tau)$ of degree at most d such that $p(\tau) \geq 0$ for all $\tau \in \mathbb{R}$. Prove that $K_+(-\infty, +\infty)$ is a closed convex cone with a compact base and that the polynomials that span extreme rays of $K_+(-\infty, +\infty)$ are

$$p(\tau) = \delta \prod_{i=1}^{k} (\tau - \tau_i)^2, \quad 2k \leq d,$$

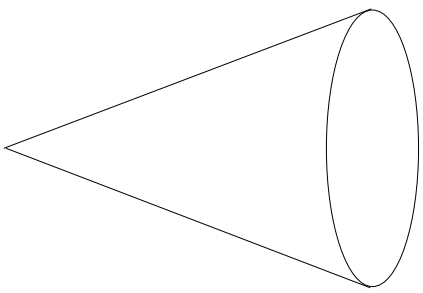
where $\delta > 0$.

Deduce that every polynomial p which is non-negative on $(-\infty, +\infty)$ can be represented in the form

$$p(\tau) = \sum_{i \in I} q_i^2(\tau),$$

where q_i are polynomials with all real roots.

4. Let $d = 2$. Draw a picture of the cone of the quadratic polynomials $a\tau^2 + b\tau + c$ in $\mathbb{R}^3 = \{(a, b, c) : a, b, c \in \mathbb{R}\}$ that are non-negative on $(-\infty, +\infty)$ in the axes a, b and c, describe the sections of the cone by the planes $a = 0$, $b = 0$ and $c = 0$ and find a compact base of the cone.



The cone of non-negative polynomials $at^2 + bt + c$ (a general view)

Figure 21

5*. Prove that every polynomial $p(\tau)$ of degree $2k$ such that $p(\tau) \geq 0$ for all $\tau \in \mathbb{R}$, admits a unique representation of the form

$$p(\tau) = \delta \prod_{i=1}^{k} (\tau - \tau_i)^2 + \gamma \prod_{i=1}^{k-1} (\tau - \tau_i')^2,$$

where $\delta, \gamma > 0$ and $\tau_1 < \tau_1' < \tau_2 < \tau_2' < \ldots < \tau_{k-1} < \tau_{k-1}' < \tau_k$, see S. Karlin and W.J. Studden, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Wiley, Interscience, 1966.

In particular, every non-negative polynomial is the sum of only two squares of polynomials.

19. The Cone of Positive Semidefinite Matrices

(19.1) The space of symmetric matrices. An $n \times n$ matrix $A = (\alpha_{ij})$ is called *symmetric*, provided $\alpha_{ij} = \alpha_{ji}$ for $i, j = 1, \ldots, n$. We identify the vector space Sym_n of all $n \times n$ symmetric matrices A with the Euclidean space $\mathbb{R}^d$, where $d = n(n+1)/2$. Let A be an $n \times n$ symmetric matrix and let U be an $n \times n$ orthogonal matrix (that is, $U^t = U^{-1}$). Then $U^{-1}AU$ is a symmetric matrix. For every symmetric $n \times n$ matrix A there is an orthogonal matrix U such that $U^{-1}AU$ is a diagonal matrix, having the eigenvalues of A on the diagonal. The number of non-zero eigenvalues is equal to the rank of A.

The scalar product of two symmetric matrices $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ is defined as

$$\langle A, B \rangle = \sum_{i,j=1}^{n} \alpha_{ij}\beta_{ij}.$$

An important formula for the scalar product is

$$\langle A, B \rangle = \mathrm{tr}(AB) = \mathrm{tr}(BA),$$

where tr is the trace, that is the sum of all diagonal entries. In particular, it follows that if U is $n \times n$ orthogonal matrix, then

$$\langle A, B \rangle = \langle U^{-1}AU, U^{-1}BU \rangle,$$

since $\langle U^{-1}AU, U^{-1}BU \rangle = \mathrm{tr}(U^{-1}AUU^{-1}BU) = \mathrm{tr}(U^{-1}ABU) = \mathrm{tr}(AB) = \langle A, B \rangle$ (we use that $\mathrm{tr}(C) = \mathrm{tr}(U^{-1}CU)$ for any invertible U). With a symmetric matrix $A = (\alpha_{ij})$ we associate the quadratic form $q_A : \mathbb{R}^n \longrightarrow \mathbb{R}$:

$$q_A(x) = \sum_{i,j=1}^{n} \alpha_{ij}\xi_i\xi_j,$$

where $x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$. If $x = (\xi_1, \ldots, \xi_n)$, we denote by $x \otimes x$ the $n \times n$ symmetric matrix $X = (x_{ij})$ where $x_{ij} = \xi_i\xi_j$. Then one can write

$$q_A(x) = \langle A, x \otimes x \rangle.$$

PROBLEMS.

1°. Prove that the dimension of the space of symmetric $n \times n$ matrices is indeed $n(n+1)/2$.

2°. Prove that $\sum_{i,j=1}^{n} \alpha_{ij} = \beta_{ij} = \mathrm{tr}(AB)$, where $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ are symmetric $n \times n$ matrices.

3°. Check that U^tAU is a symmetric matrix, provided U is any $n \times n$ matrix and A is an $n \times n$ symmetric matrix.

(19.2) Positive semidefinite matrices. An $n \times n$ symmetric matrix A is called *positive semidefinite* provided $q_A(x) \geq 0$ for all $x \in \mathbb{R}^n$. An $n \times n$ symmetric matrix A is called *positive definite* provided A is positive semidefinite and $q_A(x) = 0$ only if $x = 0$. The set of all positive semidefinite $n \times n$ symmetric matrices we denote $\mathcal{S}_+$.

Some linear algebra. A symmetric matrix is positive semidefinite if and only if all eigenvalues of A are non-negative, and positive definite if and only if all eigenvalues are positive. In particular, if A is a positive (semi)definite matrix and U is an orthogonal matrix, then $U^{-1}AU$ is a positive (semi)definite matrix. In other words, for any orthogonal matrix U, the linear transformation $X \longmapsto U^{-1}XU$ of Sym_n maps the set $\mathcal{S}_+$ onto itself. If A is positive semidefinite, then all diagonal entries are non-negative and $a_{ij}^2 \leq a_{ii}a_{jj}$ for every pair $1 \leq i \neq j \leq n$. Sylvester's criterion states that a symmetric $n \times n$ matrix A is positive semidefinite if and only if the determinant of each upper left corner submatrix $(a_{ij}) : i, j = 1, \ldots, k$ is non-negative.

PROBLEMS.

1°. Prove that $\mathcal{S}_+ \subset \mathrm{Sym}_n$ is a closed convex cone.

2. Prove that A is an interior point of $\mathcal{S}_+$ if and only if A is positive definite.

3. Prove that for every two points $x, y \in \operatorname{int} \mathcal{S}_+$ there exists a non-degenerate linear transformation T of Sym_n, such that $T\big(\mathcal{S}_+\big) = \mathcal{S}_+$ and $T(x) = y$. In other words, the cone $\mathcal{S}_+$ is *homogeneous*. Prove that the cone $\mathbb{R}^d_+ = \Big\{(\xi_1, \ldots, \xi_d) : \xi_i \geq 0$ for $i = 1, \ldots, d\Big\}$ is also homogeneous.

4. Prove that $B = \big\{A \in \mathcal{S}_+ : \operatorname{tr}(A) = 1\big\}$ is a compact base of $\mathcal{S}_+$.

5. The space Sym_2 is identified with $\mathbb{R}^3$. Draw a picture of the cone of positive semidefinite 2×2 matrices.

6°. Check, that if A is a positive semidefinite 2×2 matrix then $a_{11}, a_{22} \geq 0$ and $a_{12}^2 \leq a_{11}a_{22}$.

Now is the main result of this section.

(19.3) Proposition. *Let A be an $n \times n$ positive semidefinite matrix. Suppose that* $\operatorname{rank} A = r$. *If $r = n$, then A is an interior point of $\mathcal{S}_+$. If $r < n$, then A is an interior point of a face $\mathcal{F}$ of $\mathcal{S}_+$, where* $\dim \mathcal{F} = r(r+1)/2$. *There is a rank-preserving isometry identifying face $\mathcal{F}$ with the cone of positive semidefinite $r \times r$ matrices.*

Proof. If $\operatorname{rank} A = n$, then A is positive definite, so the result follows by Problem 2, Section 19.2.

Suppose that $\operatorname{rank} A = r < n$. We will construct a hyperplane $H \subset \mathrm{Sym}_n$ which contains A and isolates $\mathcal{S}_+$. Let $\lambda_1, \ldots, \lambda_r > 0$ be the non-zero eigenvalues of A. Let us find an orthogonal matrix U, such that $U^{-1}AU = D$, where D is the diagonal matrix such that $\lambda_1, \ldots, \lambda_r$ are the first r diagonal entries of D and all other entries

of D are 0, $D = \operatorname{diag}\big(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0\big)$. Let $C = \operatorname{diag}\big(0, \ldots, 0, 1, \ldots, 1\big)$ be matrix whose first r diagonal entries are 0 and the last $n - r$ diagonal entries are 1's, and let $Q = UCU^{-1}$.

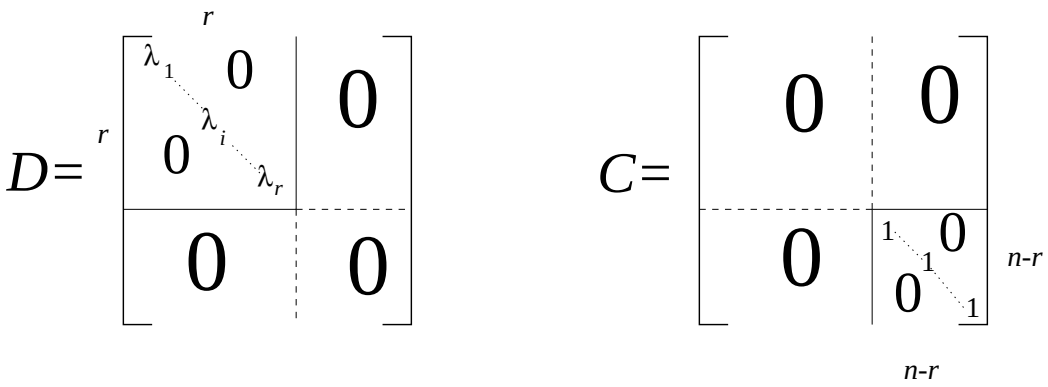


Figure 22

Then Q is a non-zero positive semidefinite matrix and

$$\langle Q, A\rangle = \langle UCU^{-1}, UDU^{-1}\rangle = \langle C, D\rangle = 0.$$

Furthermore, for any positive semidefinite $n \times n$ matrix X, the matrix $Y = U^{-1}XU$ is positive semidefinite and

$$\langle Q, X\rangle = \langle UCU^{-1}, UYU^{-1}\rangle = \langle C, Y\rangle \geq 0,$$

since the diagonal entries of Y must be non-negative. Therefore, the hyperplane $H = \big\{X \in \operatorname{Sym}_n : \langle Q, X\rangle = 0\big\}$ isolates $\mathcal{S}_+$ and contains A. Let us describe the corresponding face

$$\mathcal{F} = \big\{X \in \mathcal{S}_+ : \langle Q, X\rangle = 0\big\}.$$

The map $X \longmapsto Y = U^{-1}XU$ is a non-degenerate linear transformation which maps $\mathcal{S}_+$ onto itself, maps Q onto C and A onto D. Then the face $\mathcal{F}$ is mapped onto a face $\mathcal{F}'$, containing D and consisting of all positive semidefinite matrices Y such that $\langle Y, C\rangle = 0$:

$$\mathcal{F}' = \big\{Y \in \mathcal{S}_+ : \langle Y, C\rangle = 0\big\}.$$

Clearly, Y must have the last $n - r$ diagonal entries equal to zero. Since Y positive semidefinite, all entries in the last $n - r$ rows and last $n - r$ columns must be 0 (see Section 19.2). The upper left $r \times r$ submatrix of Y can be an arbitrary positive

semidefinite matrix.

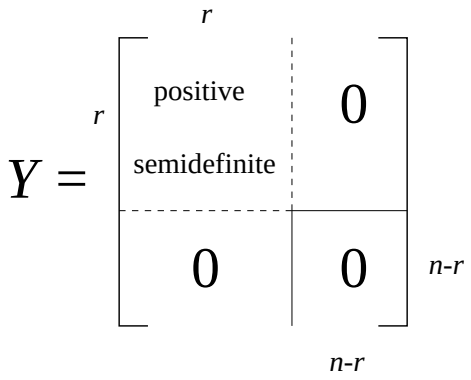


Figure 23

Thus the face $\mathcal{F}'$ may be identified with the cone of all $r \times r$ positive semidefinite matrices (in particular, $\dim \mathcal{F}' = (r+1)r/2$) and it is seen that $\mathcal{F}'$ contains D in its interior. Since $Y \longmapsto X = UYU^{-1}$ is a non-degenerate linear transformation, which maps D onto A and $\mathcal{F}'$ onto $\mathcal{F}$, we conclude that $\dim \mathcal{F} = r(r+1)/2$ and $\mathcal{F}$ contains A in its interior. □

PROBLEMS.

1. Prove that the dimensions of faces $\mathcal{F}$ of $\mathcal{S}_+$ are $0, 1, 3, \ldots, r(r+1)/2, \ldots$. Prove that if $\mathcal{F}$ is a face of $\mathcal{S}$ and $\dim \mathcal{F} = r(r+1)/2$, then there is a matrix $A \in \operatorname{int} \mathcal{F}$ such that $\operatorname{rank} A = r$.

2. Let $\mathcal{S}_+$ be the cone of $n \times n$ positive semidefinite matrices, let $\mathcal{F} \subset \mathcal{S}_+$ be a face and let r be a positive integer such that $\dim \mathcal{F} < r(r+1)/2 \leq n(n-1)/2$. Prove that there is a face $\mathcal{F}'$ of $\mathcal{S}_+$ such that $\mathcal{F}$ is a face of $\mathcal{F}'$ and $\dim \mathcal{F}' = r(r+1)/2$.

3. Let us choose positive integers $0 < r < n$, let $\mathcal{S}_1$ be the cone of positive semidefinite $n \times n$ matrices and let $\mathcal{S}_2$ be the cone of positive semidefinite $r \times r$ matrices. Let $\mathcal{F} \subset \mathcal{S}_1$ be a face such that $\dim \mathcal{F} = r(r+1)/2$. Construct an isometry (that is, a distance-preserving map) $\mathcal{S}_2 \longrightarrow \mathcal{F}$.

4. Prove that $A \in \mathcal{S}_+$ spans an extreme ray of $\mathcal{S}_+$ if and only if $\operatorname{rank} A = 1$.

Using Proposition 19.3, we get the following nice description of the facial structure of the cone of positive semidefinite matrices $\mathcal{S}_+$.

(19.4) Corollary. *The faces of $\mathcal{S}_+ \subset \operatorname{Sym}_n$ are parameterized by the subspaces of $\mathbb{R}^n$. For a (linear) subspace $L \subset \mathbb{R}^n$, let*

$$\mathcal{F}_L = \Big\{ Y \in \mathcal{S}_+ : L \subset \ker Y \Big\}.$$

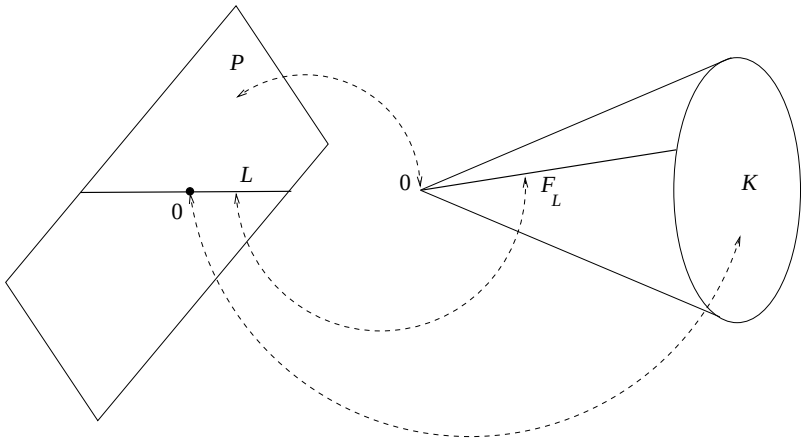
Then $\mathcal{F}_L$ is a face of $\mathcal{S}_+$ and $\dim \mathcal{F}_L = r(r+1)/2$, where $r = \operatorname{codim} L$. As L ranges over all subspaces of codimension r, $\mathcal{F}_L$ ranges over all faces of dimension $r(r+1)/2$.

Proof. Given a subspace L of codimension r, let us choose the coordinates so that $L = \{(0, \dots, 0, \xi_{r+1}, \dots, \xi_n)\}$. Then $\mathcal{F}_L$ consists of the matrices Y depicted on Figure 23. The supporting hyperplane for $\mathcal{F}_L$ is $H = \{X : \langle C, X\rangle = 0\}$, where C is depicted on Figure 22. If $\mathcal{F}$ is a face of $\mathcal{S}_+$ then $\mathcal{F} = \mathcal{F}_L$, where $L = \ker A$ and A is a matrix in the interior of $\mathcal{F}$. □

PROBLEM.

1. Let L_1 and L_2 be subspaces of $\mathbb{R}^n$. Prove that $\mathcal{F}_{L_1}$ is a face of $\mathcal{F}_{L_2}$ if and only if $L_2 \subset L_1$.

Problem 1 asserts that the *face lattice* of the cone of $n \times n$ positive semidefinite matrices is (anti)isomorphic to the lattice of all subspaces of $\mathbb{R}^n$.



Correspondence between subspaces and faces for *n=2*

Figure 24

20. Linear Equations in Positive Semidefinite Matrices

Often, to express that X is positive semidefinite, instead of writing $X \in \mathcal{S}_+$, we write $X \succeq 0$. To express that X is positive definite, we write $X \succ 0$. Proposition 19.3 has an important implication: if a system of linear equations in positive semidefinite matrices has a solution, it has a solution of a small rank.

(20.1) Proposition. *Let $\mathcal{A} \subset \mathrm{Sym}_n$ be an affine subspace such that the intersection $\mathcal{S}_+ \cap \mathcal{A}$ is non-empty and $\operatorname{codim} \mathcal{A} < (r+2)(r+1)/2$ for some non-negative integer r. Then there is a matrix $X \in \mathcal{S}_+ \cap \mathcal{A}$ such that $\operatorname{rank} X \le r$.*

Equivalently, let us fix k symmetric $n \times n$ matrices $A_1, \dots, A_k$ and k real numbers $\alpha_1, \dots, \alpha_k$. If there is a matrix $X \succeq 0$ such that

$$\langle A_i, X\rangle = \alpha_i \quad \textit{for} \quad i = 1, \dots, k,$$

then there is a matrix $X_0 \succeq 0$ such that

$$\langle A_i, X_0\rangle = \alpha_i \quad \textit{for} \quad i = 1, \dots, k$$

and, additionally

$$\operatorname{rank} X_0 \leq \left\lfloor \frac{\sqrt{8k+1}-1}{2} \right\rfloor.$$

Proof. To see that the two statements are indeed equivalent, let

$$\mathcal{A} = \Big\{ X \in \operatorname{Sym}_n : \langle A_i, X \rangle = \alpha_i : i = 1, \ldots, k \Big\}$$

be the affine subspace of symmetric $n \times n$ matrices which satisfy the given k matrix equations. Then $\operatorname{codim} \mathcal{A} \leq k$ and $k < (r+2)(r+1)/2$ if and only if $r \leq \big\lfloor \big(\sqrt{8k+1} - 1\big)/2 \big\rfloor$.

We prove the first statement. Let $\mathcal{K} = \mathcal{S}_+ \cap \mathcal{A}$. The set $\mathcal{K}$ is non-empty and closed. First, we prove that $\mathcal{K}$ has an extreme point. Let us consider a linear function $\ell(X) = \operatorname{tr}(X)$ on $\mathcal{K}$. Let $Y \in \mathcal{K}$ be a point and let $\beta = \operatorname{tr}(Y)$. We note that the set $\big\{ X \succeq 0 : \operatorname{tr}(X) \leq \beta \big\}$ is non-empty and compact, since every entry of a positive semidefinite matrix X with the trace β must not exceed β, see Section 19.2. Therefore, $\ell(X)$ attains its minimum α_0 on $\mathcal{K}$ and $\mathcal{G} = \big\{ X \in \mathcal{K} : \ell(X) = \alpha_0 \big\}$ is a compact face of $\mathcal{K}$. By the Krein-Milman Theorem (Theorem 10.3), $\mathcal{G}$ must have an extreme point X_0, which by Part 2 of Theorem 10.2 must be an extreme point of $\mathcal{K}$.

Suppose that $\operatorname{rank} X_0 = m$. Then, by Proposition 19.3, X_0 must be an interior point of a face $\mathcal{F}$ of $\mathcal{S}_+$ of dimension $m(m+1)/2$. We observe that X_0 is an interior point of the intersection $\mathcal{F} \cap \mathcal{A}$. Since X_0 is an extreme point, we must have $\dim(\mathcal{F} \cap L) = 0$, which implies that $\operatorname{codim} \mathcal{A} > \dim \mathcal{F}$, so $\operatorname{codim} \mathcal{A} > m(m+1)/2$. Hence $m \leq r$ and the proof follows. □

Some immediate consequences for systems of two matrix equations:

(20.2) Corollary. *Let us fix two symmetric $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ and two real numbers α and β. The system of two quadratic equations*

$$\sum_{i,j=1}^{n} a_{ij} \xi_i \xi_j = \alpha \quad \text{and} \quad \sum_{i,j=1}^{n} b_{ij} \xi_i \xi_j = \beta$$

has a real solution $x = (\xi_1, \ldots, \xi_n)$ if and only if the system of two linear matrix equations

$$\langle A, X \rangle = \alpha \quad \text{and} \quad \langle B, X \rangle = \beta$$

has a positive semidefinite solution $X \succeq 0$.

Proof. Let $x = (\xi_1, \ldots, \xi_n)$ be a solution to the system of quadratic equations. Let us define $X = (x_{ij})$ by $(x_{ij} = \xi_i \xi_j)$. Then X is a positive semidefinite matrix and $\langle A, X \rangle = \alpha$ and $\langle B, X \rangle = \beta$.

On the other hand, suppose there is a solution $X \succeq 0$ to the system of equations $\langle A, X \rangle = \alpha$ and $\langle B, X \rangle = \beta$. Proposition 20.1 implies that there is a solution such that $\operatorname{rank} X \leq 1$ (substitute $k = 2$ in the formula). Then $X = (x_{ij})$ can be written as $x_{ij} = \xi_i \xi_j$ for some set of numbers $\xi_1, \ldots, \xi_n$. Then $x = (\xi_1, \ldots, \xi_n)$ is a solution to the system of quadratic equations. □

(20.3) Corollary. *Let $q_1, q_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms. Consider the map $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^2$, $\phi(x) = \big(q_1(x), q_2(x)\big)$. Then the image $\phi(\mathbb{R}^n)$ is a convex cone in $\mathbb{R}^2$.*

Proof. Let $A = (a_{ij})$ be the matrix of q_1 and let $B = (b_{ij})$ be the matrix of q_2, so $q_1(x) = \sum_{i,j=1}^n a_{ij}\xi_i\xi_j$ and $q_2(x) = \sum_{i,j=1}^n b_{ij}\xi_i\xi_j$ for $x = (\xi_1, \ldots, \xi_n)$. By Corollary 20.2, the image $\phi(\mathbb{R}^n)$ can be viewed as the image of the convex cone $\mathcal{S}_+$ of positive semidefinite matrices under the linear transformation $X \longmapsto \big(\langle A, X\rangle, \langle B, X\rangle\big)$ and hence is a convex cone. □

PROBLEMS.

1°. Construct an example of a system of three quadratic equations

$$\sum_{i,j=1}^n a_{ij}\xi_i\xi_j = \alpha, \quad \sum_{i,j=1}^n b_{ij}\xi_i\xi_= \beta, \quad \sum_{i,j=1}^n c_{ij}\xi_i\xi_j = \gamma$$

such that it does not have a solution $(\xi_1, \ldots, \xi_n)$ but the corresponding system of linear matrix equations

$$\langle A, X\rangle = \alpha, \quad \langle B, X\rangle = \beta, \quad \langle C, X\rangle = \gamma$$

has a positive semidefinite solution $X \succeq 0$.

2°. Check that $X = (x_{ij})$ has the form $x_{ij} = \xi_i\xi_j$ if and only if $\operatorname{rank} X \leq 1$ and X is positive semidefinite.

4. Let $A_1, \ldots, A_k$ be $n \times n$ symmetric matrices and let $\alpha_1, \ldots, \alpha_k$ be real numbers. Let

$$\mathcal{K} = \Big\{X \succeq 0 : \langle A_i, X\rangle = \alpha_i,\ i = 1, \ldots, k\Big\}.$$

Suppose that $X \in \mathcal{K}$ and that $\operatorname{rank} X = r$. Let us decompose $X = QQ^t$, where Q is an $n \times r$ matrix of rank r. Prove that the dimension of the smallest face of $\mathcal{K}$ containing X is equal to the codimension of $\operatorname{span}\big(Q^tA_1Q, \ldots, Q^tA_kQ\big)$ in the space of $r \times r$ symmetric matrices.

For this and related results, see M.M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Springer, 1997.

5. Prove the following strengthening of Corollary 20.2: Let us fix two symmetric $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ and three real numbers α, β and γ. The system of two quadratic equations

$$\sum_{i,j=1}^n a_{ij}\xi_i\xi_j = \alpha \quad \text{and} \quad \sum_{i,j=1}^n b_{ij}\xi_i\xi_j = \beta$$

has a real solution $x = (\xi_1, \ldots, \xi_n)$ such that $\sum_{i=1}^n \xi_i^2 \leq \gamma$ if and only if the system of two linear matrix equations

$$\langle A, X\rangle = \alpha \quad \text{and} \quad \langle B, X\rangle = \beta$$

has a positive semidefinite solution $X \succeq 0$ such that $\mathrm{tr}(X) \leq \gamma$.

6. Prove the following strengthening of Corollary 20.3: Let $q_1, q_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms. Consider the map $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^2$, $\phi(x) = \big(q_1(x), q_2(x)\big)$. Then the image $\phi(B)$ of the unit ball $B = \big\{x : \|x\| \leq 1\big\}$ is a convex compact set in $\mathbb{R}^2$.

In general, as we will see shortly, the bound of Proposition 20.1 is the best possible. However, there is one important special case where it can be sharpened.

(20.4) Proposition. *Let $\mathcal{A} \subset \mathrm{Sym}_n$ be an affine subspace such that the intersection $\mathcal{S}_+ \cap \mathcal{A}$ is non-empty and bounded. Suppose that $\mathrm{codim}\,\mathcal{A} = (r+2)(r+1)/2$ for some positive $r > 0$ and that $n \geq r+2$. Then there is a matrix $X \in \mathcal{S}_+ \cap \mathcal{A}$ such that $\mathrm{rank}\, X \leq r$.*

Equivalently, for some $r > 0$, let us fix $k = (r+2)(r+1)/2$ symmetric $n \times n$ matrices $A_1, \ldots, A_k$ where $n \geq r+2$ and k real numbers $\alpha_1, \ldots, \alpha_k$. If there is a solution $X \succeq 0$ to the system

$$\langle A_i, X \rangle = \alpha_i \quad \textit{for} \quad i = 1, \ldots, k$$

and the set of all such solutions is bounded, then there is a matrix $X_0 \succeq 0$ such that

$$\langle A_i, X_0 \rangle = \alpha_i \quad \textit{for} \quad i = 1, \ldots, k$$

and, additionally

$$\mathrm{rank}\, X_0 \leq r.$$

The proof requires some algebraic topology. Below, we explicitly state what we need.

(20.5) Topological Fact. *Let us consider the set $\mathbb{RP}^{n-1}$ of all straight lines l in $\mathbb{R}^n$ passing through the origin. We make $\mathbb{RP}^{n-1}$ a metric space by letting the distance $d(l_1, l_2)$ between two lines to be the angle between l_1 and l_2. Let $\mathbb{S}^{n-1} = \big\{x : \|x\| = 1\big\}$ be the unit sphere in $\mathbb{R}^n$. Then, for $n > 2$ there is no continuous map $\phi : \mathbb{S}^{n-1} \longrightarrow \mathbb{RP}^{n-1}$ such that $\phi(x) \neq \phi(y)$ for every pair of distinct points $x, y \in \mathbb{S}^{n-1}$.*

The space $\mathbb{RP}^{n-1}$ is called the *projective space*. It is an $(n-1)$-dimensional compact connected manifold without boundary. The fact follows from the observation that $\mathbb{RP}^{n-1}$ and $\mathbb{S}^{n-1}$ are not homeomorphic for $n > 2$ and from the Invariance Domain Theorem which implies that such an embedding $\phi : \mathbb{S}^{n-1} \longrightarrow \mathbb{RP}^{n-1}$ would have been a homeomorphism, see, for example W.S. Massey, *Singular Homology Theory*, Graduate Texts in Mathematics, vol. 70, Springer-Verlag, New York, 1980.

PROBLEM.

1. Construct a continuous map $\phi : \mathbb{S}^1 \longrightarrow \mathbb{RP}^1$ such that $\phi(x) \neq \phi(y)$ for every pair of distinct points $x, y \in \mathbb{S}^1$. That is, $\mathbb{S}^1$ and $\mathbb{RP}^1$ *are* homeomorphic.

Proposition 20.4 will be deduced from the following special case.

(20.6) Lemma. *Let $r > 0$ and let $\mathcal{A} \subset \mathrm{Sym}_{r+2}$ be an affine subspace such that $\dim \mathcal{A} = r + 2$ and so $\operatorname{codim} \mathcal{A} = (r + 2)(r + 1)/2$. Suppose that the intersection $\mathcal{A} \cap \mathcal{S}_+$ is non-empty and bounded. Then there is a matrix $X \in \mathcal{S}_+ \cap \mathcal{A}$ such that $\operatorname{rank} X \leq r$.*

Proof. Suppose that $\mathcal{A} \cap \operatorname{int} \mathcal{S}_+ = \emptyset$. Problem 1 of Section 9.9 implies that $\mathcal{A}$ lies in the support hyperplane of a proper face $\mathcal{F}$ of $\mathcal{S}_+$. By Proposition 19.3, the face $\mathcal{F}$ can be identified with the cone of positive semidefinite matrices of a smaller rank $s < r + 2$, see Figure 23. Now the result follows from Proposition 20.1

Hence we may assume that $\mathcal{A} \cap \operatorname{int} \mathcal{S}_+ \neq \emptyset$. Let $\mathcal{B} = \mathcal{A} \cap \mathcal{S}_+$. Then $\mathcal{B}$ is an $(r + 2)$-dimensional convex compact set. We are going to prove that for some matrix $X_0 \in \partial\mathcal{B}$ one has $\operatorname{rank} X_0 \leq r$. Let us suppose this is not so and obtain a contradiction. For every $X \in \partial\mathcal{B}$ we must have $\operatorname{rank} X < r + 2$ (cf. Problem 2 of Section 19.2). Assuming that $\operatorname{rank} X > r$, we conclude that we must have $\operatorname{rank} X = r + 1$ for every $X \in \partial\mathcal{B}$. Therefore, for every $X \in \partial\mathcal{B}$, the set $\ker X$ is a straight line passing through the origin in $\mathbb{R}^{r+2}$. Let us construct a map $\phi : \mathbb{S}^{r+1} \longrightarrow \mathbb{RP}^{r+1}$ as follows. We consider $\mathbb{S}^{r+1}$ to be centered at a point $o \in \operatorname{int} \mathcal{B}$. For every $y \in \mathbb{S}^{r+1}$, the ray $[o, y)$ intersects $\partial\mathcal{B}$ at a single matrix $X(y)$ (we use that $\mathcal{B}$ is compact and Lemma 9.2). Let $\phi(y) = \ker X(y)$.

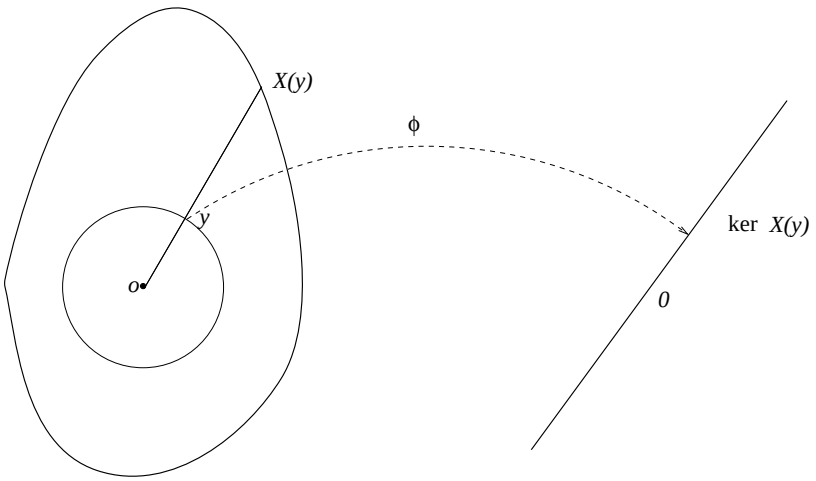


Figure 25

One can observe that ϕ is a continuous map. Since $r > 0$, by using (20.5), we conclude that there must be two distinct points $y, z \in \mathbb{S}^{r+1}$ such that $\phi(x) = \phi(y)$. In other words, there will be two distinct matrices Y and Z in $\partial\mathcal{B}$ such that $\ker Y = \ker Z$. Using Corollary 19.4, we conclude that Y and Z lie on the same face $\mathcal{F}_L$ of $\mathcal{S}_+$, where $L \subset \mathbb{R}^{r+1}$ is a straight line. Hence, by Proposition 19.3 and Corollary

19.4, the interior of $\mathcal{F}_L$ consists of matrices of rank $r+1$.

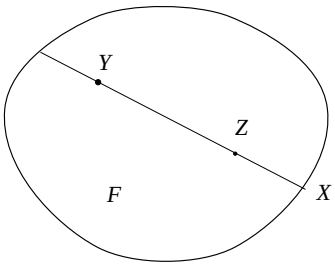


Figure 26

Let us draw the straight line (YZ) through Y and Z. Since $\mathcal{A}$ is an affine subspace, $(YZ) \subset \mathcal{A}$. Since $\mathcal{B}$ is bounded, the line (YZ) intersects $\partial\mathcal{F}$ at some point X. We must have $\operatorname{rank} X \leq r$. Clearly, $X \in \mathcal{A} \cap \mathcal{S}_+$. □

Now we are ready to prove Proposition 20.4.

Proof of Proposition 20.4. By Proposition 20.1, there is a matrix $Y \in \mathcal{A} \cap \mathcal{S}_+$ such that $\operatorname{rank} Y \leq r+1$. Let us choose a linear subspace $L \subset \mathbb{R}^n$ such that $L \subset \ker Y$ and $\operatorname{codim} L = r+2$ and let $\mathcal{F}_L$ be the corresponding face of $\mathcal{S}_+$ (see Corollary 19.4). Hence $Y \in \mathcal{F}_L$ and, therefore, $\mathcal{F}_L \cap \mathcal{A} \neq \emptyset$. Since there is a rank-preserving isometry between $\mathcal{F}_L$ and the cone of positive semidefinite $(r+2) \times (r+2)$ matrices, the proof follows by Lemma 20.6. □

PROBLEMS.

1. Show by examples that in Proposition 20.4, none of the conditions: $\mathcal{S}_+ \cap \mathcal{A}$ is bounded, $r > 0$ and $n \geq r+2$ can be dropped.

2°. Let $q_k : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms with matrices $A_k = (a_{ij,k})$, $k = 1, \ldots, m$, so

$$q_k(x) = \sum_{i,j=1}^{n} a_{ij,k}\xi_i\xi_j \quad \text{for} \quad x = (\xi_1, \ldots, \xi_n).$$

Let α_k: $k = 1, \ldots, m$ be real numbers. Suppose that the system of quadratic equations $q_k(x) = \alpha_k$ for $k = 1, \ldots, m$ has a solution $x \in \mathbb{R}^n$. Prove that there exists a positive semidefinite $n \times n$ matrix X such that $\langle A_k, X \rangle = \alpha_k$ for $k = 1, \ldots, m$ and such that $\operatorname{rank} X \leq 1$.

3. Let $A = (a_{ij}), B = (b_{ij})$ and $C = (c_{ij})$ be $n \times n$ symmetric matrices, where $n \geq 3$, and let α, β and γ be real numbers. Suppose that the system of linear matrix equations

$$\langle A, X \rangle = \alpha, \langle B, X \rangle = \beta \quad \text{and} \quad \langle C, X \rangle = \gamma$$

has a positive semidefinite solution $X \succeq 0$. Suppose further, that for some numbers τ_1, τ_2 and τ_3 the linear combination $\tau_1 A + \tau_2 B + \tau_3 C$ is (strictly) positive definite. Prove that there exists a solution $x \in \mathbb{R}^n$ to the system of quadratic equations

$$q_1(x) = \alpha, \quad q_2(x) = \beta \quad \text{and} \quad q_3(x) = \gamma,$$

where

$$q_1(x) = \sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j, \quad q_2(x) = \sum_{i,j=1}^{n} b_{ij}\xi_i\xi_j \quad \text{and} \quad q_3(x) = \sum_{i,j=1}^{n} c_{ij}\xi_i\xi_j$$

$$\text{for} \quad x = (\xi_1, \ldots, \xi_n)$$

are the corresponding quadratic forms.

21. Applications: Quadratic Convexity Theorems

In this section, we survey various *hidden convexity* results regarding the image of a quadratic map. As the main tool, we use Proposition 20.4.

(21.1) Theorem. *Let $n > 2$ and let $\mathbb{S}^{n-1} = \big\{x \in \mathbb{R}^n : \|x\| = 1\big\}$ be the unit sphere. Let $q_1, q_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms and let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^2$ be the corresponding quadratic map, $\phi(x) = \big(q_1(x), q_2(x)\big)$. Then the image $\phi(\mathbb{S}^{n-1}) \subset \mathbb{R}^2$ is a convex set.*

Proof. Let $A = (a_{ij})$ be the matrix of q_1 and let $B = (b_{ij})$ be the matrix of q_2, so $q_1(x) = \sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j$ and $q_2(x) = \sum_{i,j=1}^{n} b_{ij}\xi_i\xi_j$ for $x = (\xi_1, \ldots, \xi_n)$. Let $\mathcal{B} = \big\{X \succeq 0 : \operatorname{tr}(X) = 1\big\}$. Clearly, $\mathcal{B} \subset \operatorname{Sym}_n$ is a convex set. Let us define a linear transformation $\psi : \operatorname{Sym}_n \longrightarrow \mathbb{R}^2$ by $\psi(X) = \big(\langle A, X\rangle, \langle B, X\rangle\big)$. Clearly, $\psi(\mathcal{B}) \subset \mathbb{R}^2$ is a convex set.

We claim that $\phi(\mathbb{S}^{n-1}) = \psi(\mathcal{B})$. Indeed, if $(\alpha, \beta) \in \phi(\mathbb{S}^{n-1})$ then there is a vector $x \in \mathbb{R}^n$ such that $q_1(x) = \alpha$, $q_2(x) = \beta$ and $\|x\| = 1$. Let us define $X = (x_{ij})$ by $x_{ij} = \xi_i\xi_j$ for $i, j = 1, \ldots, n$. Then $X \in \mathcal{B}$ and $\psi(X) = (\alpha, \beta)$. Conversely, let $(\alpha, \beta) \in \psi(\mathcal{B})$ be a point. Then there exists an $X \succeq 0$ such that

$$\langle A, X\rangle = \alpha, \quad \langle B, X\rangle = \beta \quad \text{and} \quad \operatorname{tr}(X) = 1. \tag{21.1.1}$$

We observe that the set of solutions $X \succeq 0$ to the system (21.1.1) of 3 linear equations is non-empty and bounded (cf. Problem 4, Section 19.2). Applying Proposition 20.4 with $r = 1$, we conclude that there is a matrix $X_0 \succeq 0$ satisfying (21.1.1) and such that $\operatorname{rank} X_0 \leq 1$. Such a matrix can be written in the form $X_0 = (x_{ij})$, $x_{ij} = \xi_i\xi_j$ for some vector $x = (\xi_1, \ldots, \xi_n)$. We have $\|x\| = \operatorname{tr}(X_0) = 1$, $q_1(x) = \alpha$ and $q_2(x) = \beta$. Hence $(\alpha, \beta) \in \phi(\mathbb{S}^{n-1})$. □

The first proof of Theorem 21.1 was found by L. Brickman in 1961. Since then, many different proofs appeared. One of the simplest proofs is sketched below in Problem 6.

PROBLEMS.

1°. Show that Theorem 21.1 does not hold for $n = 2$.

2. Show that for any $n \geq 1$ one can find three quadratic forms $q_1, q_2, q_3 : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that if $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^3$ is the corresponding quadratic map $x \longmapsto \big(q_1(x), q_2(x), q_3(x)\big)$ then the image $\phi(\mathbb{S}^{n-1}) \subset \mathbb{R}^3$ is not convex.

3. Deduce Corollary 20.3 from Theorem 21.1.

4. Deduce the result of Problem 6, Section 20.3 from Theorem 21.1.

5. Let $q_1, q_2, q_3 : \mathbb{R}^n \longrightarrow \mathbb{R}$, $n \geq 3$, be quadratic forms and let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^3$ be the corresponding quadratic map, $\phi(x) = \big(q_1(x), q_2(x), q_3(x)\big)$. Suppose that for some numbers α_1, α_2 and α_3, the form $q = \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_3$ is (strictly) positive definite. Prove that the image $\phi(\mathbb{R}^n)$ is a convex cone in $\mathbb{R}^3$.

6. Find a different proof of Theorem 21.1 along the following lines: First, show that it would suffice to prove Theorem 21.1 for $n = 3$. Next, observe that to prove that $\phi(\mathbb{S}^2)$ is convex, it suffices to prove that the intersection of $\phi(\mathbb{S}^2)$ with every straight line $l \subset \mathbb{R}^2$, $l = \big\{(\xi, \eta) : \alpha\xi + \beta\eta = \gamma\big\}$ is connected. Now, let $q = \alpha q_1 + \beta q_2$ and prove that $\phi(\mathbb{S}^2) \cap l$ is the image of the set $\big\{x \in \mathbb{S}^2 : q(x) = \gamma\big\}$, which consists of at most 2 connected components (circles) symmetric about the origin.

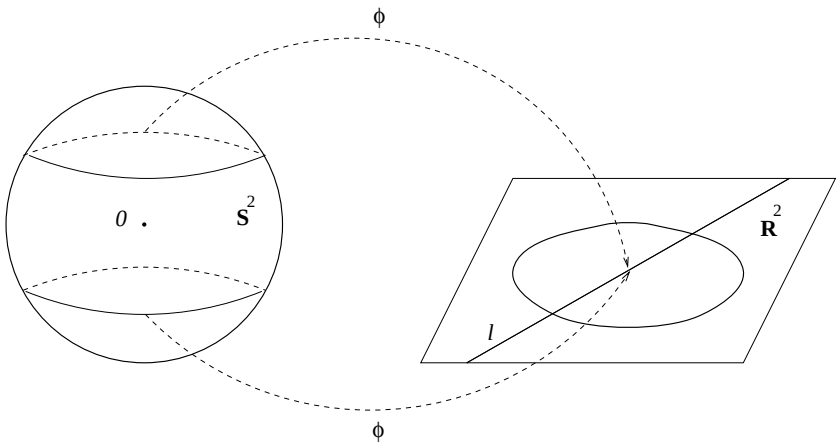


Figure 27

Theorem 21.1 implies a classical result about convexity of the numerical range of a linear operator. Given an $n \times n$ complex matrix $A = (a_{ij})$, the set $R(A) \subset \mathbb{C}$ in the complex plane:

$$R(A) = \Big\{ \sum_{i,j=1}^{n} a_{ij}\zeta_i\overline{\zeta}_j : |\zeta_1|^2 + \ldots + |\zeta_n|^2 = 1 \Big\}$$

is called the *numerical range* of A.

(21.2) Corollary (Toeplitz-Hausdorff Theorem). *The numerical range of a matrix is a convex set in the complex plane.*

Proof. For $n = 1$ the numerical range is just a point $a_{11} \in \mathbb{C}$. For $n > 1$, the set of n-tuples $(\zeta_1, \ldots, \zeta_n)$ of complex numbers $\zeta_k = \xi_k + i\eta_k$, $k = 1, \ldots, n$ such that $|\zeta_1|^2 + \ldots + |\zeta_n|^2 = 1$ can be identified with the $(2n-1)$-dimensional sphere

$$\mathbb{S}^{2n-1} = \Big\{(\xi_1, \eta_1, \xi_2, \eta_2, \ldots, \xi_n, \eta_n) : \sum_{k=1}^{n} \xi_k^2 + \sum_{k=1}^{n} \eta_k^2 = 1\Big\}.$$

The proof now follows from Theorem 21.1 since the numerical range can be viewed as the image of $\mathbb{S}^{2n-1}$ under a quadratic map $\phi : \mathbb{S}^{2n-1} \longrightarrow \mathbb{R}^2 = \mathbb{C}$. $\square$

A theory parallel to that of Sections 19-20 for real symmetric matrices, can be developed for complex Hermitian matrices and even for quaternionic Hermitian matrices. The corresponding results in the complex case are sketched in the problems below.

PROBLEMS.

1°. An $n \times n$ complex matrix $A = (a_{ij})$ is called *Hermitian* provided $a_{ij} = \overline{a_{ji}}$ for all $1 \leq i, j \leq n$. Prove that all $n \times n$ Hermitian matrices form a real vector space Her_n of dimension n^2 with the scalar product $\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij}\overline{b_{ij}}$. Prove that $\langle U^*AU, U^*BU \rangle = \langle A, B \rangle$ for any two Hermitian matrices A and B and any unitary matrix U, where * denotes the conjugate matrix.

2. A Hermitian matrix $A = (a_{ij})$ is called *positive semidefinite* provided $\sum_{i,j=1}^{n} a_{ij}\zeta_i\overline{\zeta_j} \geq 0$ for all n-tuples $z = (\zeta_1, \ldots, \zeta_n)$ of complex numbers. Let $\mathcal{H}_+ \subset \mathrm{Her}_n$ be the set of all positive semidefinite Hermitian $n \times n$ matrices. Prove that $\mathcal{H}_+$ is a closed convex n^2-dimensional cone with a compact base consisting of positive semidefinite matrices of trace 1. Draw a picture of the base of the cone for $n = 2$ (it is a 3-dimensional object).

3. Let A be an $n \times n$ positive semidefinite Hermitian matrix. Suppose that $\operatorname{rank} A = r$. Prove that A is an interior point of a face $\mathcal{F}$ of $\mathcal{H}_+$, where $\dim \mathcal{F} = r^2$. Prove that there is a rank preserving isometry identifying face $\mathcal{F}$ with the cone of positive semidefinite $r \times r$ Hermitian matrices. Prove that the faces of $\mathcal{H}_+$ are parameterized by complex linear subspaces $L \subset \mathbb{C}^n$:

$$\mathcal{F}_L = \big\{A \in \mathcal{H}_+ : L \subset \ker A\big\}.$$

4. Let us fix a number $r \geq 0$, a number $k \leq r^2 + 2r$ and a number $n \geq r$. Let $A_1, \ldots, A_k$ be $n \times n$ Hermitian matrices and let $\alpha_1, \ldots, \alpha_k$ be real numbers. Suppose that there is a positive semidefinite solution $X \in \mathrm{Her}_n$ to the system of equations $\langle A_i, X \rangle = \alpha_i$ for $i = 1, \ldots, k$. Prove that there is a positive semidefinite solution X_0 to the above system such that $\operatorname{rank} X_0 \leq r$.

5. Let us fix a number $r \geq 1$, let $k = (r+1)^2$ and let $n \geq r+2$. Let $A_1, \ldots, A_k$ be $n \times n$ Hermitian matrices and let $\alpha_1, \ldots, \alpha_k$ be real numbers. Suppose that there is a positive semidefinite solution $X \in \mathrm{Her}_n$ to the system of equations $\langle A_i, X \rangle = \alpha_i$ for $i = 1, \ldots, k$ and that the set of all such solutions is bounded. Prove that there is a positive semidefinite solution X_0 to the above system such that $\operatorname{rank} X_0 \leq r$.

6. Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$ be $n \times n$ complex Hermitian matrices, where $n \geq 3$ and let $q_1, q_2, q_3 : \mathbb{C}^n \longrightarrow \mathbb{R}$ be the corresponding Hermitian forms:

$$q_1(z) = \sum_{i,j=1}^{n} a_{ij}\zeta_i\overline{\zeta_j}, \quad q_2(z) = \sum_{i,j=1}^{n} b_{ij}\zeta_i\overline{\zeta_j} \quad \text{and} \quad q_3(z) = \sum_{i,j=1}^{n} c_{ij}\zeta_i\overline{\zeta_j}$$

for $\quad z = (\zeta_1, \ldots, \zeta_n)$.

Let $\phi : \mathbb{C}^n \longrightarrow \mathbb{R}^3$ be the map $\phi(z) = \big(q_1(z), q_2(z), q_3(z)\big)$ and let

$$\mathbb{S}^{2n-1} = \Big\{(\zeta_1, \ldots, \zeta_n) : \sum_{i=1}^{n} |\zeta_i|^2 = 1\Big\}$$

be the unit sphere. Prove that the image $\phi\big(\mathbb{S}^{2n-1}\big)$ is a convex set in $\mathbb{R}^3$.

Finally, let us describe the convex hull of a general quadratic image of the sphere. The following result was obtained by Y.H. Au-Yeung and Y.T. Poon in 1979.

(21.3) Theorem. *Let us fix a number $r \geq 1$, a number $k < (r+2)(r+1)/2$ and a number $n \geq r+2$. Let $q_1, \ldots, q_k : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms and let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be the corresponding quadratic map, $\phi(x) = \big(q_1(x), \ldots, q_k(x)\big)$. Let $\mathbb{S}^{n-1} = \big\{x \in \mathbb{R}^n : \|x\| = 1\big\}$ be the unit sphere. Then every point of $\operatorname{conv}\big(\phi(\mathbb{S}^{n-1})\big)$ can be represented as a convex combination of r (not necessarily distinct) points from $\phi(\mathbb{S}^{n-1})$.*

Proof. The proof is parallel to that of Theorem 21.1. Let $y = (\eta_1, \ldots, \eta_k)$ be a point from the convex hull of $\phi(\mathbb{S}^{n-1})$. Hence we can write $y = \alpha_1\phi(x_1) + \ldots + \alpha_m\phi(x_m)$ for some points $x_1, \ldots, x_m \in \mathbb{S}^{n-1}$ and some non-negative α_i such that $\sum_{i=1}^m \alpha_i = 1$. Let A_i be the matrix of q_i for $i = 1, \ldots, k$ and let $X = \sum_{i=1}^m \alpha_i (x_i \otimes x_i)$, where $x \otimes x$ is the matrix $(\xi_i\xi_j)$ for $x = (\xi_1, \ldots, \xi_n)$. Then $X \succeq 0$ and

$$\langle A_i, X\rangle = \eta_i \quad \text{for} \quad i = 1, \ldots, k \quad \text{and} \quad \operatorname{tr}(X) = 1. \tag{21.3.1}$$

Hence the set of all positive semidefinite matrices X satisfying (21.3.1) is non-empty and bounded. Applying Proposition 20.4 we conclude that there exists a solution $X_0 \succeq 0$ of (21.3.1) such that $\operatorname{rank} X_0 \leq r$. Such a matrix X_0 can be decomposed as $X_0 = \sum_{i=1}^r \beta_i (u_i \otimes u_i)$, where $u_i \in \mathbb{S}^{n-1}$, β_i are non-negative numbers for $i = 1, \ldots, r$ and $\sum_{i=1}^r \beta_i = 1$. It follows from (21.3.1) that

$$y = \sum_{i=1}^{r} \beta_i \phi(u_i)$$

and the proof follows. □

PROBLEMS.

1°. Deduce Theorem 21.1 from Theorem 21.3.

2. Let $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$, $D = (d_{ij})$ and $E = (e_{ij})$ be $n \times n$ real symmetric matrices. Consider the quadratic map $\phi : \mathbb{C} \longrightarrow \mathbb{R}^5$, where

$$(\zeta_1, \ldots, \zeta_n) \longmapsto \Big(\sum_{i,j=1}^{n} a_{ij}\zeta_i\overline{\zeta_j}, \sum_{i,j=1}^{n} b_{ij}\zeta_i\overline{\zeta_j}, \sum_{i,j=1}^{n} c_{ij}\zeta_i\overline{\zeta_j}, \sum_{i,j=1}^{n} d_{ij}\zeta_i\overline{\zeta_j}, \sum_{i,j=1}^{n} e_{ij}\zeta_i\overline{\zeta_j}\Big).$$

Let

$$\mathbb{S}^{2n-1} = \Big\{(\zeta_1, \dots, \zeta_n) \in \mathbb{C}^n : \sum_{k=1}^{n} |\zeta_k|^2 = 1\Big\}$$

be the sphere. Prove that the image $\phi(\mathbb{S}^{2n-1})$ is a convex set in $\mathbb{R}^5$.

3. Let us fix a number $r \geq 1$. Let $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ be the unit sphere, $n \geq r+2$. Let us fix a Borel measure μ on $\mathbb{S}^{n-1}$ such that $\mu(\mathbb{S}^{n-1}) < \infty$ and a subspace L in the space of quadratic forms $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\dim L \leq (r+1)(r+2)/2 - 1$. Prove that there exist r points $x_1, \dots, x_r \in \mathbb{S}^{n-1}$ and r non-negative numbers $\lambda_1, \dots, \lambda_r$ such that

$$\int_{\mathbb{S}^{n-1}} f \, d\mu = \sum_{i=1}^{r} \lambda_i f(x_i) \quad \text{for any} \quad f \in L.$$

4°. Let $q_1, \dots, q_k : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms whose matrices are diagonal. Let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be the corresponding quadratic map. Prove that $\phi(\mathbb{S}^{n-1})$ is a convex set in $\mathbb{R}^k$.

5*. Let us call a symmetric matrix $A = (a_{ij})$ *r-diagonal* if $a_{ij} = 0$ unless $|i-j| < r$. Let $q_1, \dots, q_k : \mathbb{R}^n \longrightarrow \mathbb{R}$ be quadratic forms whose matrices are r-diagonal matrices and let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be the corresponding quadratic map, $\phi(x) = \big(q_1(x), \dots, q_k(x)\big)$. Let $B = \big\{x \in \mathbb{R}^n : \|x\| \leq 1\big\}$ be the unit ball in $\mathbb{R}^n$. Prove that every point from $\mathrm{conv}\big(\phi(B)\big)$ is a convex combination of some r points from $\phi(B)$.

6*. In Problem 4 above, is it true that every point from $\mathrm{conv}\big(\phi(\mathbb{S}^{n-1})\big)$ is a convex combination of some r points from $\phi(\mathbb{S}^{n-1})$ if n is sufficiently large?

7. Let $q_1, \dots, q_k : \mathbb{C}^n \longrightarrow \mathbb{R}$ be Hermitian forms whose matrices are real and 2-diagonal. Let $\phi : \mathbb{C}^n \longrightarrow \mathbb{R}^k$ be the corresponding quadratic map and let $B = \Big\{(\zeta_1, \dots, \zeta_n) : \sum_{i=1}^{n} |\zeta_i|^2 \leq 1\Big\}$ be the unit ball in $\mathbb{C}^n$. Deduce from Problem 4 that the image $\phi(B)$ is a convex set in $\mathbb{R}^k$.

8*. Prove the following result of S. Friedland and R. Loewy (1976), essentially equivalent to Proposition 20.4: Suppose that $2 \leq r \leq n-1$. Let $L \subset \mathrm{Sym}_n$ be a subspace in the space of symmetric matrices such that $\dim L \geq (r-1)(2n-r+2)/2$. Then L contains a non-zero matrix whose largest eigenvalue is at least of multiplicity r.

9*. Prove the following result of F. Bohnenblust (unpublished), essentially equivalent to Proposition 20.4. Suppose that $r > 0$ and that $n \geq r+2$. Let L be a subspace in the space of quadratic forms $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\dim L < (r+2)(r+1)/2 - 1$. Suppose that the following conditions is satisfied: whenever for some vectors $x_1, \dots, x_r \in \mathbb{R}^n$ one has

$$\sum_{i=1}^{r} q(x_i) = 0 \quad \text{for all} \quad q \in L,$$

one must have $x_i = 0$ for $i = 1, \dots, r$. Then L contains a positive definite form.

22. Applications: Problems of Graph Realizability

(22.1) Some Linear Algebra: Gram matrices. Let $v_1, \ldots, v_n$ be some vectors in $\mathbb{R}^d$. Let us define an $n \times n$ matrix $X = (x_{ij})$ by $x_{ij} = \langle v_i, v_j \rangle$ (we consider the usual scalar product in $\mathbb{R}^d$). Matrix X is called the *Gram matrix* of vectors $v_1, \ldots, v_n$. It is known that X is positive semidefinite and that $\operatorname{rank} X \leq d$; in fact, $\operatorname{rank} X$ is the dimension of $\operatorname{span}(v_1, \ldots, v_n)$. Conversely, if X is a positive definite matrix such that $\operatorname{rank} X \leq d$, then X is the Gram matrix of some n vectors $v_1, \ldots, v_n$ in $\mathbb{R}^d$: $x_{ij} = \langle v_i, v_j \rangle$.

(22.2) The graph realization problem. Suppose we are given an (undirected) weighted graph $G = (V, E; \rho)$, where $V = \{v_1, \ldots, v_n\}$ is the set of vertices, E is the set of edges and $\rho : E \longrightarrow \mathbb{R}_+$ is a function, which assigns to every edge $(i, j) \in E$ a non-negative number ("length") ρ_{ij}. We say that G is *d-realizable*, if one can place the vertices $v_1, \ldots, v_n$ in $\mathbb{R}^d$ in such a way that $\|v_i - v_j\| = \rho_{ij}$ for any edge $(i, j) \in E$. We say that G is *realizable*, if it is d-realizable for some d. The most intuitive case is that 3-realizability: some problems of robotics ("linkages") and computational chemistry ("molecules") lead to problems of 3-realizability of certain graphs, see also G.M. Crippen and T.F. Havel, *Distance Geometry and Molecular Conformation*, Wiley, New York, 1988. Questions of d-realizability for the smallest possible d turn out to be relevant to problems in statistics, archaeology, genetics and geography.

PROBLEMS.

Some classical problems of the field:

1^*. Prove that one can place 7 points $v_1, \ldots, v_7$ in $\mathbb{R}^3$ in such a way that $\|v_1 - v_2\| = \|v_2 - v_3\| = \ldots = \|v_6 - v_7\| = \|v_7 - v_1\| = 1$ and $\|v_1 - v_3\| = \|v_2 - v_4\| = \ldots = \|v_5 - v_7\| = \|v_1 - v_6\| = \|v_2 - v_7\| = \sqrt{8/3}$.

The constants 1 and $\sqrt{8/3}$ are chosen in such a way that the angles between consecutive intervals v_{i-1}, v_i and v_i, v_{i+1} are equal to $\arccos -1/3$, that is, to the angle between two intervals connecting vertices of a regular tetrahedron with its

center.

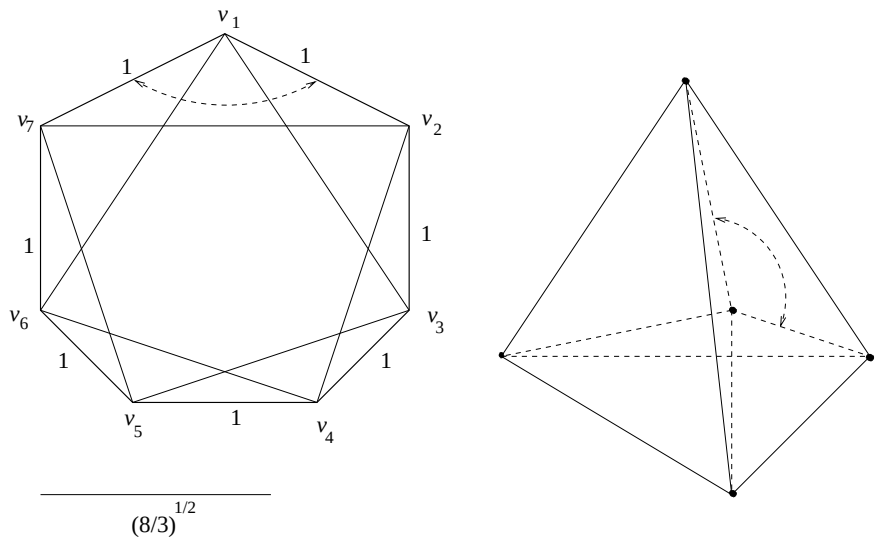


Figure 28

2*. Prove that every such a configuration of 7 points in $\mathbb{R}^3$ has one degree of freedom (modulo rigid motions).

3*. Prove that one can place 6 points $v_1, \ldots, v_6$ in $\mathbb{R}^3$ in such a way that $\|v_1 - v_2\| = \|v_2 - v_3\| = \ldots = \|v_5 - v_6\| = \|v_6 - v_1\| = 1$ and $\|v_1 - v_3\| = \|v_2 - v_4\| = \ldots = \|v_4 - v_6\| = \|v_1 - v_5\| = \|v_2 - v_6\| = \sqrt{8/3}$.

4*. Prove that there are two such configurations of 6 points, one of which has one degree of freedom (modulo rigid motions) and the other is rigid.

5* Consider the X set of all configurations of 5 points $v_1, \ldots, v_5$ in $\mathbb{R}^2$ such that $\|v_1 - v_2\| = \|v_2 - v_3\| = \|v_3 - v_4\| = \|v_4 - v_5\| = \|v_5 - v_1\| = 1$. Thus X can be viewed as a subset of $\left(\mathbb{R}^2\right)^5 = \mathbb{R}^{10}$. We observe that if $x \in X$ and g is an orientation preserving isometry of $\mathbb{R}^2$ then $g(x) \in X$. Let Y be the factor space of X modulo all orientation preserving isometries of $\mathbb{R}^2$. Prove that Y is a 2-dimensional manifold homeomorphic to the sphere with 4 handles.

An easy problem.

6°. Prove that if a weighted graph with n vertices is realizable, it is $(n-1)$-realizable.

(22.3) A straightforward reformulation. Let $v_1, \ldots, v_n$ be a realization of the graph in $\mathbb{R}^d$, and let $X = (x_{ij})$, $x_{ij} = \langle v_i, v_j \rangle$ be the Gram matrix of $v_1, \ldots, v_n$. Then $X \succeq 0$ and for any edge $e = (i, j)$ of G, we have

$$\rho_{ij} = \|v_i - v_j\|^2 = \langle v_i, v_i \rangle - 2\langle v_i, v_j \rangle + \langle v_j, v_j \rangle = x_{ii} - 2x_{ij} + x_{jj}$$

Hence we conclude that the problem of realizability of G is equivalent to the following problem:

Is there an $n \times n$ matrix $X = (x_{ij})$ such that $X \succeq 0$ and

$$x_{ii} - 2x_{ij} + x_{jj} = \rho_{ij}^2$$

for every edge $(i, j) \in E$?

The problem of d-realizability is equivalent to the above problem with one additional constraint:

$$\operatorname{rank} X \leq d.$$

It turns out, that if d is large enough, realizability is equivalent to d-realizability.

(22.4) Proposition. *Let k be the number of edges of G. Suppose that $k < (d+2)(d+1)/2$. Then G is d-realizable, if and only if it is realizable. In particular, if $k \leq 9$, the graph is realizable if and only if it is 3-realizable.*

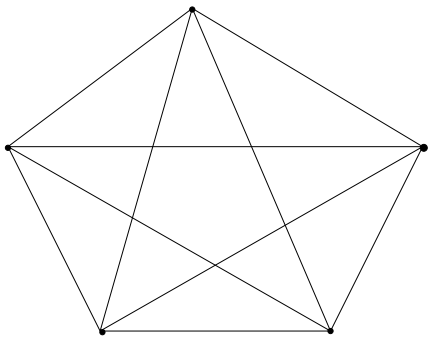
Proof. By Proposition 20.1, if the system of linear matrix equations

$$x_{ii} - 2x_{ij} + x_{jj} = \rho_{ij}^2 : \quad (i, j) \in E$$

has a positive semidefinite solution $X = (x_{ij})$, it has a positive semidefinite solution X_0, such that additionally, $\operatorname{rank} X_0 \leq d$. The proof now follows. □

PROBLEMS.

1. Let G be the complete graph with $(d+2)$ vertices (and $(d+2)(d+1)/2$ edges) such that the length of every edge (v_i, v_j) is 1. Prove that G is realizable but not d-realizable.



The complete graph with 5 vertices

Figure 29

2. Suppose that G is a cycle $v_1 - v_2 - \ldots - v_n - v_1$ with some weights on the edges. Prove that G is realizable, if and only if it is 2-realizable.

3*. Suppose that G has n vertices $v_1, \ldots, v_n$ and $2n$ edges: $(v_1, v_2), (v_2, v_3), \ldots, (v_{n-1}, v_n), (v_n, v_1)$ and $(v_1, v_3), (v_2, v_4), \ldots, (v_{n-2}, v_n), (v_{n-1}, v_1), (v_n, v_2)$

with some weights on the edges. Is it true, that if G is realizable, it is 4-realizable?

4^*. A graph $G = (V, E)$ is called *chordal*, provided for any cycle $v_1 - v_2 - \ldots - v_k - v_1$, where $k \geq 4$ there is a chord $v_i - v_j$, where $1 \leq i < j \leq k$. A subset $K \subset V$ is called a *clique*, if every two vertices from K are connected by an edge. Prove that G is chordal if and only if it has the following property:

for any choice of weights on the edges, the graph is realizable, provided every clique is realizable.

See M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Springer, 1997.

5. Suppose that we want to embed a hexagon $v_1, \ldots, v_6$ in $\mathbb{R}^d$ with prescribed distances $\|v_1 - v_2\|$, $\|v_2 - v_3\|$, $\|v_3 - v_4\|$, $\|v_4 - v_5\|$, $\|v_5 - v_6\|$ and $\|v_6 - v_1\|$ and prescribed angles between the pairs of opposite edges: (v_1, v_2) and (v_4, v_5), (v_2, v_3) and (v_5, v_6) and (v_3, v_4) and (v_6, v_1).

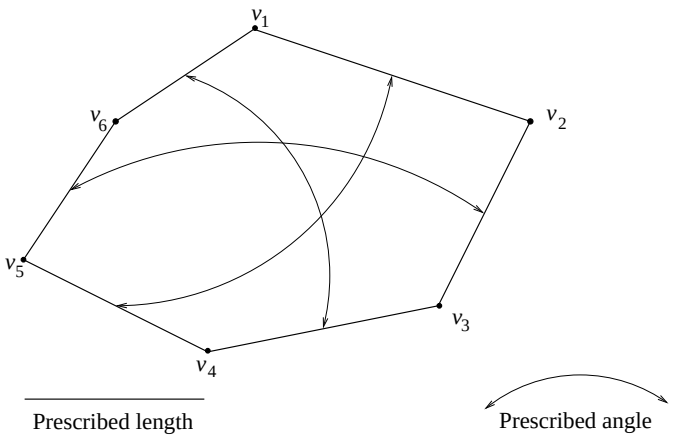


Figure 30

Prove that if such an embedding exists for some d, it exists for $d = 3$.

Problem 1 of Section 22.4 shows that the bounds of Proposition 22.4 (and hence the bounds of Proposition 20.1) are the best possible. It turns out, however, the the graph of Problem 1 is the only graph with $(d+2)(d+1)/2$ edges which is realizable, but not d-realizable.

(22.5) The economical reformulation. Let $v_1, \ldots, v_n$ be a realization of the graph $G = (V, E)$ with n vertices in $\mathbb{R}^d$. We can always assume that $v_n = 0$. Let $X = (x_{ij})$ be the $(n-1) \times (n-1)$ Gram matrix of the vectors $v_1, \ldots, v_{n-1}$. Then $X \succeq 0$ and we have the following affine constraints:

(22.5.1)
$$x_{ii} = \rho_{in}^2 \quad \text{if } (i, n) \text{ is an edge} \quad \text{and}$$
$$x_{ii} - 2x_{ij} + x_{jj} = \rho_{ij}^2 \quad \text{if } (i, j) \text{ is an edge and } 1 \leq i, j \leq n-1.$$

Hence we conclude that the problem of realizability of G is equivalent to the following problem:

Is there an $(n-1)\times(n-1)$ matrix $X=(x_{ij})$ such that $X\succeq 0$ and X satisfies (22.5.1)?

The problem of d-realizability is equivalent to the above problem with one additional constraint:

$$\operatorname{rank} X \le d.$$

(22.6) Proposition. *Let k be the number of edges of G. Suppose that $k=(d+2)(d+1)/2$ and that G is not a union of a complete graph with $d+2$ vertices and 0 or more isolated vertices. Then G is d realizable, if and only if it is realizable.*

Proof. Since G is d-realizable if and only if its connected components are realizable, without loss of generality we may assume that G is connected. Since G is not a complete graph, we must have $n\ge d+3$, and so $n-1\ge d+2$. Now we use Proposition 20.4. We claim that the set of positive semidefinite solutions X to the system 22.5.1 is bounded. Indeed, since G is connected, each vertex v_i can be connected to v_n by a path. Since v_n is fixed at 0, the length of $\|v_i\|=\sqrt{x_{ii}}$ is bounded by the length of the path. In particular, $\sqrt{x_{ii}}$ is bounded by the sum of all ρ_{ij}. Therefore, the set of feasible matrices $X\succeq 0$ is bounded (cf. Section 19.2). Hence the proof follows by Proposition 20.4. □

PROBLEMS.

1. Let G be a graph with $n\ge d+2$ vertices consisting of a complete graph with $(d+2)$ vertices and $n-d-2$ isolated vertices. Let us assign length 1 to every edge. Prove that G is realizable, but not d realizable. Letting $n=d+2$, deduce that the condition $n\ge d+2$ in Proposition 20.4 can not be dropped. Letting $n>d+2$, deduce that the condition of boundedness in Proposition 20.4 can not be dropped either.

2. Given a graph $G=(V,E)$ with $|V|=n$ vertices and $k=|E|$ edges and a number d, let us consider the following *rigidity map* $\phi_G:\mathbb{R}^{nd}\longrightarrow\mathbb{R}^k$. A point x from $\mathbb{R}^{nd}=\mathbb{R}^d\times\ldots\times\mathbb{R}^d$ is interpreted as an n-tuple of d vectors $(v_1,\ldots,v_n)$ and $\phi(x)$ is the k-tuple of squared distances $\|v_i-v_j\|^2$, where (i,j) runs over all edges of G. Let $\mathbb{S}^{nd-1}$ be the unit sphere in $\mathbb{R}^{nd}$. Prove that if $n\ge d+3$ and $k<(d+2)(d+1)/2$ then the image $\phi(\mathbb{S}^{nd-1})$ is a convex set in $\mathbb{R}^k$.

23. Separation Theorems in Euclidean Space

Let V be a vector space. Recall (see Definition 8.4) that sets $A,B\subset V$ are separated by a hyperplane H if A and B belong to different closed halfspaces $\overline{H_+}$ and $\overline{H_-}$. Equivalently, the sets A and B are separated by a hyperplane if there exist a non-zero linear function $f:V\longrightarrow\mathbb{R}$ and a number α such that $f(x)\le\alpha$ for each $x\in A$ and $f(x)\ge\alpha$ for each $x\in B$. Sets $A,B\subset V$ are strictly separated by a hyperplane H if A and B belong to different open halfspaces H_+ and H_-. Equivalently, the sets A and B are strictly separated by a hyperplane if there exist

a non-zero linear function $f : V \longrightarrow \mathbb{R}$ and a number α such that $f(x) < \alpha$ for all $x \in A$ and $f(x) > \alpha$ for all $x \in B$.

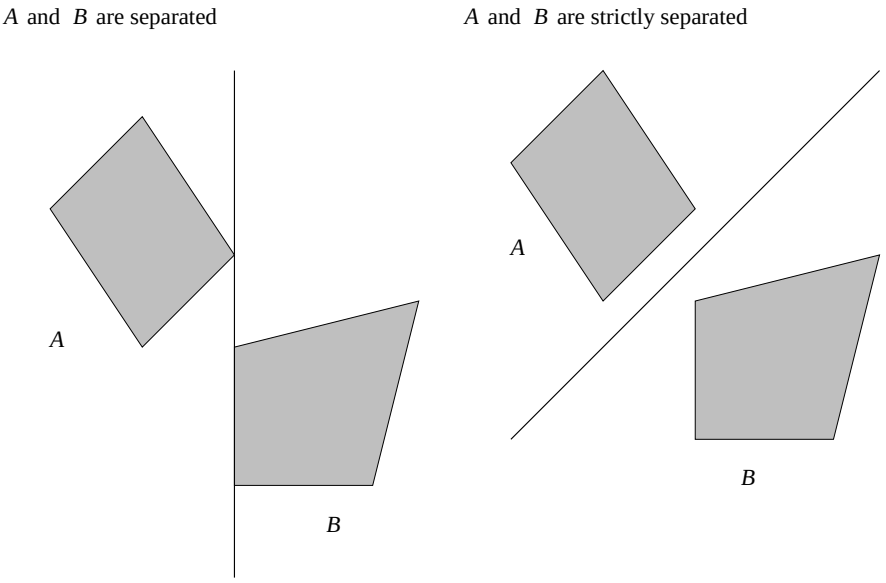


Figure 31

(23.1) Definition. Let V be a vector space and let $A, B \subset V$ be sets. The *Minkowski sum* $A + B$ is defined as

$$A + B = \{x + y : \ x \in A, \ y \in B\}.$$

PROBLEM.

1°. Prove that if sets A and B are convex, their Minkowski sum $A + B$ is convex as well.

(23.2) Theorem. *Let A and B be non-empty convex sets in $\mathbb{R}^d$ such that $A \cap B = \emptyset$. Then there exists a hyperplane H which separates A and B.*

Proof. Consider the set

$$C = A - B = \{x - y : x \in A, y \in B\}.$$

We claim that C is convex. Let $u = x_1 - y_1$ and $v = x_2 - y_2$, where $x_1, x_2 \in A$ and $y_1, y_2 \in B$. Then, for any $0 \leq \alpha \leq 1$, we have $x = \alpha x_1 + (1 - \alpha)x_2 \in A$, $y = \alpha y_1 + (1 - \alpha)y_2 \in B$ and

$$\alpha u + (1 - \alpha)v = \alpha(x_1 - y_1) + (1 - \alpha)(x_2 - y_2) = x - y \in C.$$

Since $A \cap B = \emptyset$, we have $0 \notin C$. Therefore, by Theorem 9.9, there exists a hyperplane H such that $0 \in H$ and H isolates C. In other words, there exists a

non-zero linear function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, such that $f(x-y) \le 0$ for all $x \in A$, $y \in B$. We have $f(x) \le f(y)$ for each $x \in A$ and each $y \in B$. Therefore, there exists an $\alpha \in \mathbb{R}$, such that $f(x) \le \alpha$ for all $x \in A$ and $f(y) \ge \alpha$ for all $y \in B$ (we can choose $\alpha = \sup\big\{f(x) : x \in A\big\}$, for example). The hyperplane $H = \big\{x \in \mathbb{R}^d : f(x) = \alpha\big\}$ separates A and B. $\square$

PROBLEMS.

1°. Suppose that one of the sets $A, B \subset \mathbb{R}^d$ is open. Prove that $C = A - B$ is open.

2. Construct an example of two closed sets $A, B \subset \mathbb{R}^2$ such that $C = A - B$ is not closed.

3. Let $A, B \subset \mathbb{R}^d$ be sets. Suppose, that A is closed and B is compact. Prove that the set $C = A - B$ is closed.

Hint: Let $x_n \in C$ be a sequence of points and suppose that $x_n \longrightarrow x$. We have to prove that $x \in C$. We may write $x_n = y_n - z_n$, where $y_n \in A$ and $z_n \in B$. Since B is compact, there is a subsequence z_{n_k} which converges to a point of B: $z_{n_k} \longrightarrow z \in B$. Then we must have $y_{n_k} \longrightarrow y \in A$ (since A is closed). Therefore, $x = y - z \in C$.

4. Prove that if $A, B \subset \mathbb{R}^d$ are compact sets, then $C = A - B$ is a compact set.

5°. Let $A, B \subset \mathbb{R}^d$ be open sets. Suppose that a hyperplane H separates A and B. Prove, that H strictly separates A and B.

(23.3) Theorem. *Let $A \subset \mathbb{R}^d$ be a closed convex set and let $u \notin A$ be a point. Then there exists a hyperplane H which strictly separates A and u.*

Proof. Since A is a closed set and $u \notin A$, there exists an $r > 0$ such that the open ball $B(u; r) = \big\{x : \|x - u\| < r\big\}$ does not intersect A. Let $B = B(0; r/2) = \big\{x : \|x\| < r/2\big\}$. We claim that $A + B$ and $u + B$ are open non-intersecting convex sets. Indeed,

$$A + B = \bigcup_{x \in A} B(x; r/2)$$

is a union of open sets, so it is open. Similarly, $u + B = B(u; r/2)$ is an open ball centered at u. It follows by Problem 1, Section 23.1, that $A + B$ and $u + B$ are convex. Suppose that $(A + B) \cap (B + u) \ne \emptyset$. For a point $x \in (A + B) \cap (B + u)$, we have $\|u - x\| < r/2$ and $\|y - x\| < r/2$ for some $y \in A$. Therefore, $\|u - y\| < r$, which is a contradiction. Therefore, the sets $A + B$ and $B + u$ do not intersect.

By Theorem 23.2, there is a hyperplane H that separates $A + B$ and $u + B$. Since the sets $A + B$ and $u + B$ are open, H must strictly separate the sets (see Problem 5, Section 23.2). $\square$

PROBLEMS.

1. Construct an example of two non-empty convex closed non-intersecting sets A and B in $\mathbb{R}^2$ that can not be strictly separated by a hyperplane.

2. Let $A \subset \mathbb{R}^d$ be a non-empty convex closed set and let $u \notin A$ be a point. Prove that there exists a unique point $v \in A$, such that $\|u - v\| \le \|u - x\|$ for all

$x \in A$. Furthermore, prove that the hyperplane H orthogonal to $u - v$ and passing through the point $(u + v)/2$ strictly separates A and u.

3. Using Problem 3, Section 23.2, prove that if $A, B \subset \mathbb{R}^d$ are non-empty non-intersecting convex sets, where A is closed and B is compact, then A and B can be strictly separated by a hyperplane H.

4. Let $A, B \subset \mathbb{R}^d$ be non-empty non-intersecting convex sets. Suppose that B is compact and A is closed. Prove that there is a pair of poits $u \in A$ and $v \in B$ such that $\|u - v\| \leq \|x - y\|$ for all $x \in A$ and all $y \in B$. Furthermore, prove that the hyperplane H orthogonal to $u - v$ and passing through $(u + v)/2$ strictly separates A and B.

(23.4) What can go wrong in infinite dimension?

PROBLEM.

1. Let $V = \mathbb{R}_\infty$ be the vector space of all infinite sequences $x = (\xi_1, \dots)$ of real numbers such that all but finitely many terms ξ_i are zero (see Problem 2, Section 8.6). Let $A \subset V$ be the set of all such sequences x whose last non-zero term is strictly positive and let $B = -A$ be the set of sequences whose last non-zero term is strictly negative. Prove that A and B are convex, that $A \cap B = \emptyset$ and that A and B can not be separated by a hyperplane.

We need the notion of an algebraically open set (see Definition 8.5).

(23.5) Theorem. *Let V be a vector space and let $A, B \subset V$ be non-empty convex sets such that $A \cap B = \emptyset$. Suppose that A is algebraically open. Then A and B can be separated by a hyperplane.*

Proof. Let $C = A - B$. We can write C as a union of algebraically open sets:

$$C = \bigcup_{b \in B} (A - b).$$

Therefore, C is algebraically open (see Problem 6, Section 8.5). Furthermore, C is convex and $0 \notin C$. Therefore, by Theorem 8.6 there exists a hyperplane H, such that $0 \in H$ and H isolates C. The proof is completed as in Theorem 23.2. $\square$

(23.6) Definition. Let V be a vector space and let $A \subset V$ be a convex set. We say that $v \in A$ *lies in the algebraic interior* of A if for any straight line l passing through v the point v lies in the interior of the intersection $A \cap l$. The set of all points v that lie in the algebraic interior of A is called the *algebraic interior* of A.

PROBLEM.

1. Let V be a vector space and let $A \subset V$ be a convex set. Let u_0 be a point in the algebraic interior of A. Prove, that for any point $u_1 \in A$ and any $0 \leq \alpha < 1$, the point $u_\alpha = (1 - \alpha)u_0 + \alpha u_1$ lies in the algebraic interior of A.

Hint: cf. Lemma 9.2.

2. Prove that the algebraic interior of a convex set is a convex algebraically open set.

Hint: Use Problem 1.

3. Let V be a vector space, let $A, B \subset V$ be convex sets, and let $H \subset V$ be a hyperplane. Suppose that H separates B and the algebraic interior of A. Prove that H separates A and B.

Hint: Use Problem 1.

4. Let $A \subset \mathbb{R}_\infty$ be the set of Problem 1 of Section 23.4. Prove that the algebraic interior of A is empty.

(23.7) Corollary. *Let V be a vector space and let $A, B \subset V$ be non-empty convex sets such that $A \cap B = \emptyset$. Suppose that A has a non-empty algebraic interior. Then A and B can be separated by a hyperplane.*

Proof. Let A_1 be the algebraic interior of A. Then, by Problem 2, Section 23.6, A_1 is a non-empty algebraically open convex set. By Theorem 23.5, there is a hyperplane H which separates A_1 and B. Then, by Problem 3, Section 23.6, H separates A and B. □

24. Topological Vector Spaces, Convex Sets and Hyperplanes

(24.1) Topological spaces. We recall that a *topological space* is a set X together with a family $F \subset 2^X$ of its subsets, called *open sets*, such that

$\emptyset$ and X are open sets;

intersection of any two (equivalently, of finitely many) open sets is an open set;

union of open sets is a an open set.

An open subset containing a point $x \in X$ is called a *neighborhood* of that point. A set $C \subset X$ is called *closed* if $X \setminus C$ is open. A map $\phi : X \longrightarrow Y$, where X and Y are topological spaces, is called *continuous*, if for any open set $U \subset Y$ the preimage $\phi^{-1}(U) = \big\{x \in X : \phi(x) \in U\big\}$ is an open set in X. Equivalently, ϕ is continuous, if the preimage of a closed set in Y is closed in X. If $F_1, F_2 \subset 2^X$ are two topologies on X, we say that F_2 is *stronger* (F_1 is *weaker*), if $F_1 \subset F_2$.

Let X and Y be topological spaces. The direct product $Z = X \times Y$ is made a topological space by declaring a set $U \subset Z$ open if it can be represented as a union of direct products $U_1 \times U_2$, where U_1 is an open subset of X and U_2 is an open subset of Y. Equivalently, this topology is the weakest among all topologies that make the projections $Z \longrightarrow X$, $(x, y) \longmapsto x$ and $Z \longrightarrow Y$, $(x, y) \longmapsto y$ continuous.

(24.2) Topological vector spaces. Let V be a vector space. Suppose that V is also a topological space so that the following properties hold:

For every vector $v \in V$ the set $\{v\}$ is closed;

the map $V \times V \longrightarrow V$, $(x, y) \longmapsto x + y$ is continuous. Equivalently, for every $w_1, w_2 \in V$ and every neighborhood U of $u = w_1 + w_2$, there is a neighborhood W_1 of w_1 and a neighborhood W_2 of w_2 such that $W_1 + W_2 \subset U$;

the map $\mathbb{R} \times V \longrightarrow V$, $(\alpha, x) \longmapsto \alpha x$ is continuous. Equivalently, for every $w \in V$ and every $\alpha \in \mathbb{R}$, for every neighborhood U of $u = \alpha w$, there is a neighborhood W of of w and a number $\epsilon > 0$ such that $\beta x \in U$ provided $x \in W$ and $|\alpha - \beta| < \epsilon$.

Then V is called *topological vector space*.

In particular, for any given $u \in V$, the *translation* $x \longmapsto u + x$ is a continuous transformation, and for any given $\alpha \in \mathbb{R}$, the *scaling* $x \longmapsto \alpha x$ is a continuous transformation.

PROBLEMS.

1. Let V be a topological vector space and let $x \in V$ be a vector. Prove that if $U \subset V$ is an open (closed) set, then the translation $U + x$ is an open (closed) set.

2. Let V be a topological vector space and let $\alpha \neq 0$ be a number. Prove, that if $U \subset V$ is an open (closed) set, then $\alpha U = \{\alpha x : x \in U\}$ is an open (closed) set in V.

3. Let V be a topological vector space and let $w_1 \neq w_2$ be two distinct points in V. Prove, that there exist neighborhoods W_1 of w_1 and W_2 of w_2 such that $W_1 \cap W_2 = \emptyset$.

(24.3) Definition. Let V be a topological vector space and let $u \in V$ be a point. A family of neighborhoods $\{U_i\}$ of u is called a *local base* at u, provided for every neighborhood U of u, there is a neighborhood $U_i \subset U$ from the local base.

To define a topological vector space, it suffices to define its local base at the origin (see Problem 1 below).

PROBLEMS.

1. Let V be a topological vector space and let $\{U_i\}$ be a local base at the origin. Prove that every non-empty open set $U \subset V$ is a union of translations $U_i + x_i$ of some sets U_i from the base and some $x_i \in V$.

2. A set A in a vector space V is called *balanced*, provided $\alpha x \in A$ for each $x \in A$ and each α, such that $|\alpha| \leq 1$. Prove that every neighborhood of the origin in a topological vector space contains a balanced neighborhood of the origin.

(24.4) Definitions. Let V be a topological vector space, let $A \subset V$ be a set, and let $u \in V$ be a point. We say that u lies in the interior of A, provided there is a neighborhood $U \subset A$ of u. We say that u lies in the closure of A, provided for every neighborhood U of u, we have $U \cap A \neq \emptyset$. The set of all points in the interior of A is called the *interior* of A and denoted $\mathrm{int}(A)$. The set of all points in the *closure* of A is called the *closure* of A and denoted $\mathrm{cl}(A)$.

The following result is very similar to Lemma 9.2.

(24.5) Lemma. *Let V be a topological vector space and let $A \subset V$ be a convex set. Let $u_0 \in \operatorname{int}(A)$ and $u_1 \in A$. Then, for every $0 \leq \alpha < 1$ and $u_\alpha = (1-\alpha)u_0 + \alpha u_1$, we have $u_\alpha \in \operatorname{int}(A)$.*

Proof. Let $U_0 \subset A$ be a neighborhood of u_0. Let us consider a map $T : V \longrightarrow V$:

$$x \longmapsto \frac{1}{1-\alpha}(x - u_1) + u_1.$$

Then T is continuous and $T(u_\alpha) = u_0$ (see Figure 12). Therefore, the preimage $U_\alpha = T^{-1}(U_0)$ is a neighborhood of u_α. Let us show that $U_\alpha \subset A$. Indeed, if $x \in U_\alpha$, then $T(x) = y \in A$. Solving for x, we get $x = (1-\alpha)y + \alpha u_1$. Since A is convex, the proof follows. □

(24.6) Theorem. *Let V be a topological vector space and let $A \subset V$ be a convex set. Then* $\operatorname{int}(A)$ *and* $\operatorname{cl}(A)$ *are convex sets.*

Proof. The proof that $\operatorname{int}(A)$ is a convex set follows the proof of Corollary 9.3. Let us prove that $\operatorname{cl}(A)$ is a convex set. Let $u_0, u_1 \in \operatorname{cl}(A)$ and let $u_\alpha = \alpha u_0 + (1-\alpha)u_1$ for $0 < \alpha < 1$. We know that every neighborhood of u_0 or u_1 intersects A and we must prove that every neighborhood of u_α intersects A. Let U be a neighborhood of u_α.

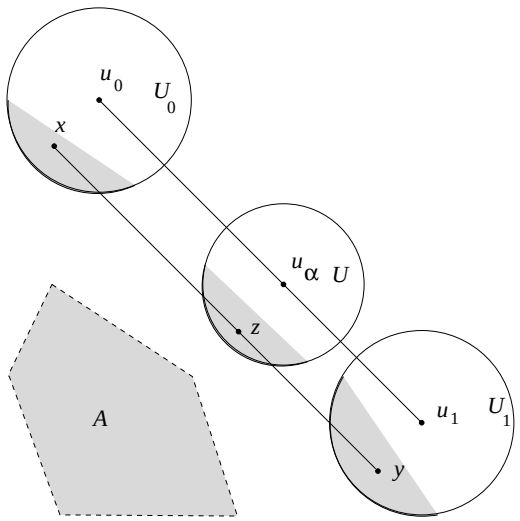


Figure 32

Since the operation $(x, y) \longmapsto \alpha x + (1-\alpha)y$ is continuous, there must be a neighborhood U_0 of u_0 and U_1 of u_1, such that $\alpha x + (1-\alpha)y \in U$ for every $x \in U_0$ and every $y \in U_1$. Since $u_0 \in \operatorname{cl}(A)$, there is a point $x \in U_0 \cap A$ and since $u_1 \in \operatorname{cl}(A)$, there is a point $y \in U_1 \cap A$. Since A is convex, for $z = \alpha x + (1-\alpha)y$ we have $z \in A$. Therefore, $U \cap A \neq \emptyset$. □

PROBLEMS.

1°. Prove that $\operatorname{int}(A+x) = \operatorname{int}(A)+x$ and $\operatorname{cl}(A+x) = \operatorname{cl}(A)+x$ for each $A \subset V$ and each $x \in V$.

2°. Prove that $\operatorname{int}(\alpha A) = \alpha \operatorname{int}(A)$ and $\operatorname{cl}(\alpha A) = \alpha \operatorname{cl}(A)$ for each $A \subset V$ and each $\alpha \neq 0$.

(24.7) Theorem. *Let V be a topological vector space and let $H \subset V$ be an affine hyperplane. Then either $\operatorname{cl}(H) = H$ (that is, H is closed), or $\operatorname{cl}(H) = V$ (that is, H is dense in V).*

Proof. It suffices to prove the result for linear hyperplanes. Let $f : V \longrightarrow \mathbb{R}$ be a linear function such that

$$H = \big\{x \in V : f(x) = 0\big\}.$$

Suppose that H is not closed. Then there is a point $u \in V$, such that $u \notin H$, but every neighborhood of u contains a point from H. Let $\alpha = f(u) \neq 0$. Let us choose any $w \in V$. Suppose that $\beta = f(w)$. Let $\gamma = \beta/\alpha$ and let $h = w - \gamma u$. Then $f(h) = f(w) - \gamma f(u) = 0$, so $h \in H$. In other words, $w = h + \gamma u$. Since multiplication and addition are continuous operations, for every neighborhood W of w, there is a neighborhood U of u, such that $h+\gamma U \subset W$. Since $u \in \operatorname{cl}(H)$, there must be a point $h_1 \in U \cap H$. Then the point $h + \gamma h_1$ lies in W. Hence $W \cap H \neq \emptyset$ and $w \in \operatorname{cl}(H)$. Thus we proved that if H is not closed, then for each point $w \in V$, we have $w \in \operatorname{cl}(H)$. □

PROBLEMS.

1. Consider the space $C[0,1]$ of all continuous functions on the interval $[0,1]$. We make $C[0,1]$ a topological vector space by choosing a local base at the origin consisting of the sets

$$U_\epsilon = \Big\{ f \in C[0,1] : |f(\tau)| < \epsilon \quad \text{for all} \quad \tau \in [0,1] \Big\}$$

for any $\epsilon > 0$. Let $L \subset C[0,1]$ be a subspace consisting of all smooth functions (a function is called smooth if its differentiable at every point and the derivative is continuous). Prove that L has infinite codimension and that $\operatorname{cl}(L) = C[0,1]$.

2*. Using Zorn's Lemma and Problem 1, show that there exists a dense hyperplane in $C[0,1]$.

3. Prove that every hyperplane in $\mathbb{R}^d$ is closed.

4. Consider the space V of all continuous functions on the interval $[0,1]$. We make V a topological vector space by choosing a local base at the origin consisting of the sets

$$U_\epsilon = \Big\{ f \in V : \int_0^1 \sqrt{|f(\tau)|}\, d\tau < \epsilon \Big\}$$

for any $\epsilon > 0$. Prove that for every hyperplane $H \subset V$, we have $\operatorname{cl}(H) = V$.

5. Let V be the space of all smooth functions (functions with continuous derivative) on the interval $[0,1]$. We make V a topological vector space by choosing a local base at the origin consisting of the sets

$$U_\epsilon = \Big\{ f \in C[0,1] : |f(\tau)| < \epsilon \quad \text{for all} \quad \tau \in [0,1] \Big\}$$

for any $\epsilon > 0$ (as in Problem 1). Let $H = \big\{ f : f'(1/2) = 0 \big\}$. Prove that H is a dense hyperplane in V.

6. Let V be a topological vector space and let $L \subset V$ be an (affine) subspace. Prove that $\operatorname{cl} L$ is an (affine) subspace.

(24.8) Theorem. *Let V be a topological vector space, let $f : V \longrightarrow \mathbb{R}$ be a nonzero linear function, and let $\alpha \in \mathbb{R}$ be a number. The affine hyperplane $H(\alpha) = \big\{ x \in V : f(x) = \alpha \big\}$ is closed if and only if f is continuous.*

Proof. If f is continuous, then $H(\alpha) = f^{-1}(\alpha)$ is a closed set, as the preimage of a closed set $\{\alpha\}$.

Let us prove that if H is closed, then f is continuous. First, let us prove that both halfspaces H_+ and H_- are open. Suppose, for example, that H_+ is not open. Then there is a point $u \in H_+$ such that every neighborhood of u intersects $\overline{H_-}$. Applying a translation, if necessary, we may assume that $u = \mathbf{0}$ is the origin. Let us choose any neighborhood U of $\mathbf{0}$. Since $0 \cdot \mathbf{0} = \mathbf{0}$ and multiplication by a scalar is continuous, there is a neighborhood W of the origin and a number $1 > \delta > 0$ such that $\tau x \in U$ for any $|\tau| < \delta$ and any $x \in W$. Let $U_1 = \bigcup_{0<\tau<\delta} \tau W$. Then $U_1 \subset U$ is a neighborhood of the origin and $[0, x] \subset U_1$ for any $x \in U_1$ (cf. Problem 2 of Section 24.3).

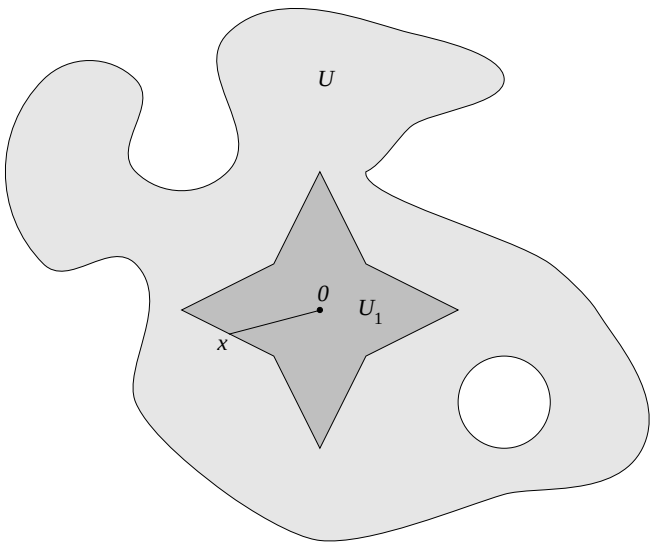


Figure 33

Hence U_1 must intersect $\overline{H_-}$. If $x \in H_-$, since f is linear, the interval $[0, x]$ intersects H and hence U_1 intersects H. If $x \in H$, then U_1 intersects H as well. Since $U_1 \subset U$, every neighborhood U of u intersects H, which contradicts the assumption that H is closed. Therefore, both H_+ and H_- are open.

Since all the halfspaces $H_-(\gamma) = \{x : f(x) < \gamma\}$ are translations of each other and all the halfspaces $H_+(\gamma) = \{x : f(x) > \gamma\}$ are translations of each other, it follows that all $H_-(\gamma)$ and all $H_+(\gamma)$ are open. Let $I = (\beta_1, \beta_2) \subset \mathbb{R}$ be an open interval. Then

$$f^{-1}(I) = \{x : \beta_1 < f(x) < \beta_2\} = H_-(\beta_2) \cap H_+(\beta_1)$$

is an open set in V. Hence every open set A in $\mathbb{R}$ is a union of open intervals, $f^{-1}(A)$ is open and f is continuous. □

(24.9) The Dual Space

Let V be a topological vector space. The dual space V^*, as a vector space, consists of all continuous linear functions $f : V \longrightarrow \mathbb{R}$ with the addition: $g = f_1 + f_2$ provided $g(x) = f_1(x) + f_2(x)$ for all $x \in V$ and scalar multiplication: $g = \alpha f$ provided $g(x) = \alpha f(x)$ for all $x \in V$. There is a remarkable topology, called weak* (pronounced "weak star") topology, on V^*. Open sets in V^* are unions of elementary open sets of the type

$$U(x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_n; \beta_1, \ldots, \beta_n)$$

$$= \Big\{ f \in V^* : \alpha_i < f(x_i) < \beta_i \quad \text{for} \quad i = 1, \ldots, n \Big\},$$

where $x_1, \ldots, x_n \in V$ and $\alpha_1, \beta_1, \ldots, \alpha_n, \beta_n \in \mathbb{R}$.

PROBLEMS.

1. Prove that for any $x \in V$, the function $\phi_x : V^* \longrightarrow \mathbb{R}$, $\phi(f) = f(x)$ is a continuous linear function on V^*.

2*. Prove that every continuous linear function $\phi : V^* \longrightarrow \mathbb{R}$ has the form $\phi(f) = f(x)$ for some $x \in V$.

3. Prove that for any two distinct points $f, g \in V^*$, there is a continuous linear function $\phi : V^* \longrightarrow \mathbb{R}$, such that $\phi(f) \neq \phi(g)$.

4*. (Banach-Alaoglu Theorem). Let $U \subset V$ be a neighborhood of the origin. Let

$$K = \{f \in V^* : \ |f(x)| \leq 1 \quad \text{for all } x \in U\}.$$

Prove that K is compact.

As a general reference for this section, see W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

25. Separation Theorems in Topological Vector Spaces

(25.1) Lemma. *Let V be a topological vector space and let $A \subset V$ be an open set. Then A is an algebraically open set.*

Proof. Let $L = \{u + \tau v : \tau \in \mathbb{R}\}$ be a straight line in V ($v \neq 0$). We must prove that the intersection $A \cap L$ is an open set in L, possibly empty. If $A \cap L$ is not empty, let $w \in A \cap L$, $w = u + \tau_0 v$. Since A is open, there is a neighborhood W of w, such that $W \subset A$. Since addition and scalar multiplication are continuous operations, there exists an $\epsilon > 0$, such that if $|\tau - \tau_0| < \epsilon$, then $u + \tau v \in W$. This implies that w is an interior point of the intersection $A \cap L$. Therefore, A is an algebraically open set. □

PROBLEMS.

1. Let $V = \mathbb{R}_\infty$ be a vector space of all infinite sequences of real numbers $x = (\xi_1, \xi_2, \dots)$, such that only finitely many terms ξ_i are non-zero (see Problem 1 of Section 23.4 and Problem 2, Section 8.6). We make V a topological vector space by choosing the local base at the origin

$$U_\epsilon = \left\{x : \sum_{i=1}^{\infty} \xi_i^2 < \epsilon\right\}$$

for $\epsilon > 0$. Let $A = \{x \in V : |\xi_k| < 1/k \text{ for } k = 1, \dots\}$. Prove that A is convex and algebraically open, but not open.

2. Let V be a vector space. Let us declare a set $A \subset V$ open if and only if it is a union of convex algebraically open sets. Prove that this converts V into a topological vector space (cf. Problem 6 of Section 8.5).

3. Let V be a topological vector space of Problem 2. Prove that every hyperplane in V is closed.

4. Let V be a topological vector space and let $A, B \subset E$ be non-empty open sets. Suppose that a hyperplane $H \subset E$ separates A and B. Prove that H strictly separates A and B and that H is closed.

(25.2) Theorem. *Let V be a topological vector space and let $A, B \subset V$ be convex sets. Suppose that $A \cap B = \emptyset$ and that $\operatorname{int}(A) \neq \emptyset$. Then there is a closed hyperplane $H \subset V$ that separates A and B. Equivalently, there is a continuous linear function $f : V \longrightarrow \mathbb{R}$ such that $f(x) \leq f(y)$ for all $x \in A$ and $y \in B$ and f is not identically zero.*

Proof. By Lemma 25.1, the algebraic interior of A is not empty. Therefore, by Corollary 23.7, there is a hyperplane $H \subset V$ that separates A and B. By Theorem 24.7, either H is closed or H is dense in V. In the latter case, H must have a non-empty intersection with any open subset, in particular, with the interior of A.

Therefore, H has a non-empty intersection with the algebraic interior of A.

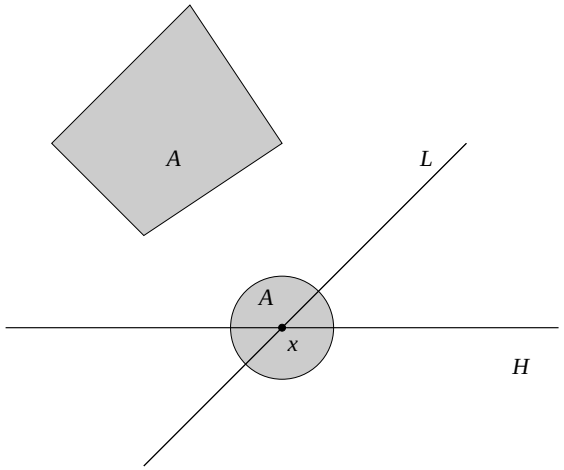


Figure 34

If $x \in H \cap A$ is a point in the algebraic interior of A, then we can choose a straight line L passing through x and not contained in H. The intersection $L \cap A$ contains an open interval around x and hence H can not isolate A and thus can not separate A and B. The contradiction shows that H must be a closed hyperplane. By Theorem 24.8, the corresponding linear function f is continuous. □

Now, a very important definition.

(25.3) Definition. A topological vector space V is called *locally convex*, provided for every point $u \in V$ and every neighborhood U of u, there is a convex neighborhood $W \subset U$ of u. In other words, there is a local base of convex neighborhoods of the origin (see Section 24.3).

PROBLEMS.

1. Let V be a vector space. A function $\|\cdot\| : E \longrightarrow \mathbb{R}$ is called a *norm*, provided

$\|u\| \geq 0$ for each $u \in E$ and $\|u\| = 0$ only if $u = 0$,

$\|\alpha u\| = |\alpha|\|u\|$ for each $u \in E$ and each $\alpha \in \mathbb{R}$,

$\|u + v\| \leq \|u\| + \|v\|$ for each $u \in E$ and each $v \in E$.

Let us make V a topological vector space by choosing a base at the origin consisting of the sets $U_\epsilon = \{x : \|x\| < \epsilon\}$ for all $\epsilon > 0$. Prove that V is a locally convex topological vector space. Such a space V is called a *normed space*.

2. Let V be a topological vector space and let $A \subset V$ be an open set. Prove that $\text{conv}(A)$ is an open set.

3. Let V be a vector space. Prove that the strongest topology that makes V a locally convex topological vector space is the topology where a set $U \subset V$ is open if and only if it is a union of convex algebraically open sets. Prove that every linear function $f : V \longrightarrow \mathbb{R}$ is continuous in this topology.

4. Let us fix p, $0 < p < 1$. Let us consider the set $L^p[0,1]$ of measurable functions such that $\int_0^1 |f(\tau)|^p \, d\tau < +\infty$ (we identify functions that coincide almost everywhere). Prove, that $L^p[0,1]$ is a vector space. Furthermore, let us choose a local base at 0, consisting of the sets

$$U_\epsilon = \left\{ f \in L^p : \int_0^1 |f(\tau)|^p \, d\tau < \epsilon \right\}$$

for all $\epsilon > 0$. Prove that this makes $L^p[0,1]$ a topological vector space. Prove, that the only convex open sets in $L^p[0,1]$ are $\emptyset$ and the whole space $L^p[0,1]$.

(25.4) Theorem. *Let V be a locally convex topological vector space. Let $A \subset V$ be a closed convex set and let $u \notin A$ be a point. Then there exists a closed hyperplane H that strictly separates A and u. Equivalently, there exists a continuous linear function $f : V \longrightarrow \mathbb{R}$ such that $f(x) < f(u)$ for all $x \in A$.*

Proof. Since A is closed, the complement $U = V \setminus A$ is a neighborhood of u. Let $U_0 = U - u$ be a translation of U, so U_0 is a neighborhood of the origin. Because the transformation $(x, y) \longmapsto x - y$ is continuous, there are neighborhoods W_1 and W_2 of the origin, such that $x - y \in U_0$ for each $x \in W_1$ and each $y \in W_2$. Since V is locally convex, we can choose W_1 and W_2 to be convex.

Consider the Minkowski sums $A + W_2$ and $u + W_1$. The sets are convex (see Problem 1, Section 23.1) and open, since $A + W_2 = \bigcup_{x \in A} (W_2 + x)$ is a union of open sets. Furthermore, $(A + W_2) \cap (u + W_1) = \emptyset$.

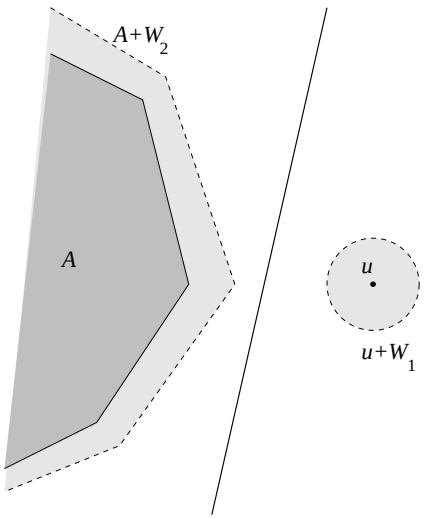


Figure 35

Indeed, if $(A+W_2)\cap(u+W_1)\neq\emptyset$, then for a point $a\in A$, a point $y\in W_2$ and a point $x\in W_1$, we will have $a+y=u+x$, that is $a=u+(x-y)$. Since $x-y\in U_0$, this would imply that A and U intersect, which is a contradiction. Theorem 25.2 implies that there is a closed hyperplane that separates $A+W_2$ and $u+W_1$. Since $A+W_1$ and $u+W_1$ are open, the hyperplane H must strictly separate $A+W_1$ and $u+W_1$ (see Problem 4 of Section 25.1). $\square$

PROBLEMS.

1°. Prove that in a locally convex topological vector space V, any two distinct points $x\neq y$ can be strictly separated by a closed hyperplane.

2. Let V be a locally convex topological vector space and let $A, B\subset V$ be convex sets, such that A is closed, B is compact, and $A\cap B=\emptyset$. Prove that A and B can be separated by a closed hyperplane.

Hint: It suffices to construct convex neighborhoods W_1 and W_2 of the origin, such that $(A+W_1)\cap(B+W_2)=\emptyset$ and then use Theorem 25.2. To achieve this goal, for every point $x\in B$, construct convex neighborhoods $W_1(x)$ and $W_2(x)$ of the origin, such that $(A+W_1(x))\cap(x+W_2(x))=\emptyset$. Since B is compact, there is a finite set of points $x_1,\ldots,x_n\in B$, such that the sets $x_i+W_2(x_i)$, $i=1,\ldots,n$ cover B. Let $W_1=\bigcap\limits_{i=1}^{n}W_1(x)$ and $W_2=\bigcap\limits_{i=1}^{n}W_2(x)$.

26. The Krein-Milman Theorem in Topological Vector Spaces

(26.1) The Krein-Milman Theorem. *Let V be a locally convex topological vector space and let $K\subset V$ be a convex compact set. Then K is the closure of the convex hull of the set of its extreme points, $K=\mathrm{cl}\Big(\mathrm{conv}\big(\mathrm{ex}(K)\big)\Big)$.*

Proof. First, we establish that every non-empty compact set K in V has an extreme point. Let us call a non-empty convex compact subset $A\subset K$ *extreme* provided for any two points $x,y\in K$ and $z=(x+y)/2$, whenever $z\in A$, we must have $x,y\in A$. Clearly, K is an extreme set. Let $X\subset K$ be the smallest extreme subset (that is, not containing any extreme subset of K, other than itself). The existence of X is established via Zorn's Lemma or the axiom of choice. Let us prove that X is a point. Suppose that X contains two different points, say x_1 and x_2. Let us choose a closed hyperplane H that strictly separates x_1 and x_2, see Theorem 25.4. In other words, there is a continuous linear function $f:V\longrightarrow\mathbb{R}$ such that $f(x_1)<f(x_2)$. Let $\alpha=\min\big\{f(y):y\in X\big\}$. Since X is compact (a closed subset of a compact) the minimum is attained. Let $Y=\big\{x\in X:f(x)=\alpha\big\}$ be a face of X. Clearly, Y is a convex compact subset of K and Y is an extreme set (cf. Theorem 10.2). On the other hand, Y does not contain x_2, so Y is strictly smaller than X. The contradiction shows that X must be a point, that is, an extreme point.

Now, we prove that K is the closure of the convex hull of the set of its extreme points. Let $A=\mathrm{cl}\Big(\mathrm{conv}\big(\mathrm{ex}(K)\big)\Big)$. Then A is a closed convex subset of K. We

must prove that $A = K$.

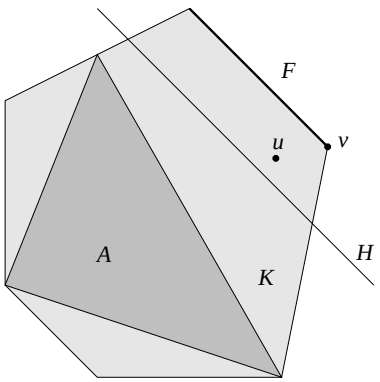


Figure 36

Suppose that there is a point $u \in K \setminus A$. Let us choose a closed hyperplane H that strictly separates u from A (Theorem 25.4). In other words, there is a continuous linear function $f : V \longrightarrow \mathbb{R}$ such that $f(x) > f(u)$ for any $x \in A$. Let $\alpha = \min\{f(x) : x \in K\}$ (the minimum is attained since f is continuous and K is compact) and let $F = \{x \in K : f(x) = \alpha\}$ be the corresponding face of K. So, F is a compact convex set and, as we proved, must contain an extreme point v, which will be an extreme point of K (cf. Theorem 10.2). On the other hand, $v \notin A$, which is a contradiction. □

PROBLEMS.

1. Let $C[0,1]$ be the vector space of all continuous functions on the interval $[0,1]$. Let $B = \{f : |f(\tau)| \leq 1 \text{ for all } x \in [0,1]\}$. Find the extreme points of B.

2. Let A be the following subset of $C[0,1]$:

$$A = \left\{f : \int_0^1 f(\tau)\, d\tau = 0 \quad \text{and} \quad |f(\tau)| \leq 1 \text{ for all } \tau \in [0,1]\right\}.$$

Prove that A has no extreme points.

3. Let V be a topological vector space and let A be a closed convex set, such that $\operatorname{int}(A) \neq \emptyset$. Prove that for each point $u \in \partial A$, there is a closed support hyperplane at u, that is, a closed hyperplane H that contains u and isolates A.

4. Let $A \subset V$ be a convex set in a vector space V and let F be a face of A. Prove, that F is an extreme set of A, that is, for any two $x, y \in A$, whenever $z = (x+y)/2 \in F$, we must have $x, y \in F$.

5. Let $L \subset C[0,1]$ be the vector space of all smooth functions on the interval $[0,1]$. We make L a topological vector space by using the $C[0,1]$ metric: $\operatorname{dist}(f,g) = \max\{|f(x) - g(x)| : x \in [0,1]\}$. Let $A = \{f : f'(1/2) > 0\}$ and $B = \{f : f'(1/2) < 0\}$ be subsets of L.

Prove that A and B are convex, non-intersecting, algebraically open subsets of L, so there is a hyperplane $H \subset L$ strictly separating them. Prove that A and B are dense in L. Deduce, that there is no closed hyperplane $H \subset L$ strictly separating A and B.

We get an infinite-dimensional counterpart of Corollary 10.4.

(26.2) Corollary. *Let V be a locally convex topological vector space and let $K \subset V$ be a convex compact set. Let $f : V \longrightarrow \mathbb{R}$ be a continuous linear function. Then there exists an extreme point u of K such that $f(u) \geq f(x)$ for all $x \in K$.*

Proof. Since f is continuous and K is compact, f attains its maximum value α on K. Let $F = \big\{x \in K : f(x) = \alpha\big\}$ be the corresponding face of K. Hence F is a compact convex set and by the Krein-Milman Theorem (Theorem 26.1), F has an extreme point u. Then by Theorem 10.2 (Part 2), u is an extreme point of K. We have $\alpha = f(u) \geq f(x)$ for all $x \in K$. □

27. Polyhedra in L^∞

(27.1) Spaces L^1 and L^∞. Let $L^1[0,1]$ be a vector space of all integrable functions on the interval $[0,1]$, that is, Lebesgue measurable functions f such that

$$\int_0^1 |f(\tau)|\, d\tau < +\infty.$$

As usual, we don't distinguish functions that differ on a set of measure 0. We make $L^1[0,1]$ a topological vector space by choosing a local base at the origin consisting of the neighborhoods

$$U_\epsilon = \Big\{f : \int_0^1 |f(\tau)|\, d\tau < \epsilon\Big\}$$

for $\epsilon > 0$. Thus a set $U \subset L^1[0,1]$ is open if and only if it is a union of translations $U_{\epsilon_i} + x_i$, see Section 24.3.

Let $L^\infty[0,1]$ be the set of all Lebesgue measurable functions f on the interval $[0,1]$ such that $|f(\tau)| \leq C$ for some constant C and for almost all (all, except from a set of zero measure) $\tau \in [0,1]$. As usual, we don't distinguish functions that differ on a set of measure 0. There is the standard (strong) topology on $L^\infty[0,1]$ defined by the local base at the origin:

$$U_\epsilon = \Big\{f : |f(\tau)| \leq \epsilon \quad \text{for almost all} \quad \tau \in [0,1]\Big\}.$$

Hence a set $U \subset L^\infty[0,1]$ is open if and only if it is a union of translations $U_{\epsilon_i} + x_i$. It is known that every continuous linear function $\phi : L^1[0,1] \longrightarrow \mathbb{R}$ has the form

$$\phi(f) = \int_0^1 f(\tau)g(\tau)\, d\tau$$

for some $g \in L^\infty[0,1]$. This allows us to introduce a weak* topology on L^∞, see Section 24.9.

(27.2) Fact. *Let $B \subset L^\infty[0,1]$ be the set*

$$B = \Big\{ f : 0 \le f(\tau) \le 1 \quad \textit{for almost all} \quad \tau \in [0,1] \Big\}.$$

The set B is compact in the weak topology.*

See, for example, W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

PROBLEMS.

1. Prove that if f is an extreme point of B then $f(\tau) \in \{0,1\}$ for almost all $\tau \in [0,1]$.

2*. Prove that the set B is compact in the weak* topology.

Hint: Consider a neighborhood U of the origin in $L^1[0,1]$:

$$U = \Big\{ g : \int_0^1 |g(\tau)|\, d\tau < 1 \Big\}.$$

Let $K = \Big\{ f \in L^\infty : \int_0^1 (fg)\, d\tau \le 1 \text{ for all } g \in U \Big\}$. Prove that K consists of the functions $f \in L^\infty[0,1]$ such that $|f(\tau)| \le 1$ for almost all $\tau \in [0,1]$. Use the Banach-Alaoglu Theorem, see Problem 4, Section 24.9, to claim that K is compact. Show that $B \subset K$ is a closed subset.

3. Prove that $\operatorname{conv}\big(\operatorname{ex}(B)\big)$ is not closed.

4. Deduce from Proposition 27.2 that the set

$$K = \Big\{ f \in L^\infty[0,1] : |f(\tau)| \le 1 \quad \text{for almost all} \quad \tau \in [0,1] \Big\}$$

is weak* compact.

Next, we introduce what may be considered as an L^∞ analogue of a polyhedron.

(27.3) Proposition. *Let us fix m functions $a_1(\tau), \dots, a_m(\tau) \in L^1[0,1]$ and m numbers $\alpha_1, \dots, \alpha_m \in \mathbb{R}$. Let $B \subset L^\infty[0,1]$ be the set*

$$B = \Big\{ f : 0 \le f(\tau) \le 1 \quad \textit{for almost all} \quad \tau \in [0,1] \Big\},$$

and let

$$A = \Big\{ f \in B : \int_0^1 a_i(\tau) f(\tau)\, d\tau = \alpha_i \quad \textit{for } i = 1, \dots, m \Big\}.$$

Then A is a convex weak compact subset of L^∞[0,1]. If u is an extreme point of A, then $u(\tau) \in \{0,1\}$ for almost all τ.*

Proof. It is obvious that A is convex. Furthermore, A is a closed subset of B and so A is compact as follows from Proposition 27.2.

Let u be an extreme point of A. Suppose that the set $\{\tau : u(\tau) \notin \{0,1\}\}$ has a positive measure. Then for some $\delta > 0$ and $X = \{\tau : \delta \le |u(\tau)| \le 1 - \delta\}$, the measure of X is positive. Let us find $m+1$ pairwise disjoint subsets $X_1, \ldots, X_{m+1}$ of X of positive measure and let $[X_i]$ be the indicator function of X_i:

$$[X_i](\tau) = \begin{cases} 1 & \text{if } \tau \in X_i \\ 0 & \text{if } \tau \notin X_i. \end{cases}$$

Let $\epsilon_1, \ldots, \epsilon_{m+1}$ be real numbers (to be specified later) and let

$$v = \epsilon_1[X_1] + \ldots + \epsilon_{m+1}[X_{m+1}].$$

Furthermore, let $u_+ = u + v$ and let $u_- = u - v$. Then $u = (u_+ + u_-)/2$. Obviously, if $|\epsilon_1| < \delta, \ldots, |\epsilon_m| < \delta$, we will have $0 \le u_+(\tau), u_-(\tau) \le 1$ for almost all τ.

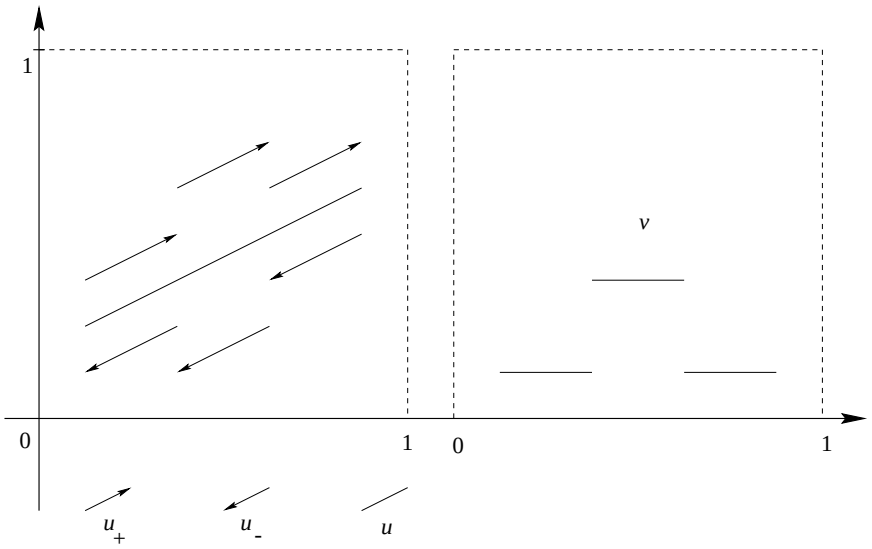


Figure 37

Now we show that we can choose a non-zero v so that $u_+, u_- \in A$. Indeed, it suffices to choose $\epsilon_1, \ldots, \epsilon_{m+1}$ in such a way, that the system of m homogeneous equations

$$\sum_{j=1}^{m+1} \epsilon_j \int_0^1 a_i[X_j]\, d\tau = 0, \quad i = 1, \ldots, m$$

in $m+1$ variables $\epsilon_1, \ldots, \epsilon_{m+1}$ is satisfied. Since the number of variables exceeds the number of equations, there must be a solution $\epsilon_1, \ldots, \epsilon_{m+1}$, where not all ϵ's are zero. Scaling, if necessary, we make $\epsilon_1, \ldots, \epsilon_{m+1} < \delta$. Then $u_-, u_+ \in A$. That contradicts the assumption that u is an extreme point of A. □

PROBLEM.

1. Let us consider the space $L^\infty([0,1], \mathbb{R}^d)$ of all Lebesgue measurable bounded vector-valued functions $[0,1] \longrightarrow \mathbb{R}^d$ and the space $L^1([0,1], \mathbb{R}^d)$ of all Lebesgue

integrable vector-valued functions $f : [0,1] \longrightarrow \mathbb{R}^d$. Let us fix a polytope $P \subset \mathbb{R}^d$, m functions $a_1, \dots, a_m \in L^1([0,1], \mathbb{R}^d)$, and m numbers $\alpha_1, \dots, \alpha_m$. Let

$$A = \Big\{ f \in L^\infty([0,1], \mathbb{R}^d) :$$

$$f(\tau) \in P \quad \text{and} \quad \int_0^1 \langle c_i, f(\tau) \rangle \, d\tau = \alpha_i \quad \text{for } i = 1, \dots, m \Big\}.$$

Prove that A is a convex set and that if f is an extreme point of A, then $f(\tau)$ is a vertex of P for almost all τ.

28. An Application: Problems of Linear Optimal Control

(28.1) Solving linear systems of differential equations with control. Let $A(\tau)$ and $B(\tau)$ be $n \times n$ matrices, where $\tau \in [0,1]$ is a real parameter. We assume that $A(\tau)$ and $B(\tau)$ are smooth functions of τ. We consider a system of linear differential equations

$$\frac{d}{d\tau}\mathbf{x}(\tau) = A(\tau)\mathbf{x}(\tau) + B(\tau)\mathbf{u}(\tau), \tag{28.1.1}$$

where $\mathbf{x}(\tau) = \big(x_1(\tau), \dots, x_n(\tau)\big)$ and $\mathbf{u}(\tau) = \big(u_1(\tau), \dots, u_n(\tau)\big)$ are vectors from $\mathbb{R}^n$. The function $\mathbf{x}(\tau)$ is a *solution* of the system (28.1.1), whereas $\mathbf{u}(\tau)$ is a *control*. If the control is chosen, under mild assumptions the solution $\mathbf{x}$ is determined by the *initial condition*:

$$\mathbf{x}(0) = \mathbf{x}_0,$$

where $\mathbf{x}_0 \in \mathbb{R}^n$ is a given vector. Under mild assumptions, the solution $\mathbf{x}$ can be found as follows. Let $X(\tau)$ be an $n \times n$ matrix, which is a solution to the matrix system

$$\frac{d}{d\tau}X(\tau) = A(\tau)X(\tau) \quad \text{with the intial condition} \quad X(0) = I,$$

where I is an $n \times n$ matrix. Then

$$\mathbf{x}(\tau) = X(\tau)\Big(\mathbf{x}_0 + \int_0^\tau X^{-1}(t)B(t)\mathbf{u}(t)\, dt\Big). \tag{28.1.2}$$

In general, we want to choose the control $\mathbf{u}$ in the space of L^∞ vector-valued functions $[0,1] \longrightarrow \mathbb{R}^n$. Then, if $X(\tau)$ and $B(\tau)$ are continuous, the integral (28.1.2) is well-defined. Let us consider $\mathbf{x}(\tau)$ as a solution to the original system of differential equations, even though it may be, say, non-differentiable at some points. In particular, if we are to choose the control $\mathbf{u}(\tau)$ in such a way that some *terminal* condition $\mathbf{x}(1) = \mathbf{x}_1$, is satisfied, we get the following integral constraint:

$$\int_0^1 X^{-1}(\tau)B(\tau)\mathbf{u}(\tau)\, d\tau = X^{-1}(1)\mathbf{x}_1 - \mathbf{x}_0. \tag{28.1.3}$$

PROBLEM.

1°. Check formula (28.1.2).

(28.2) Example. A problem of linear control. Let us fix smooth real functions $a_0(\tau), a_1(\tau), b(\tau), c_0(\tau), c_1(\tau)$, and $d(\tau)$: $\tau \in [0,1]$. Consider a differential equation for a function $x(\tau)$

$$\ddot{x}(\tau) = a_0(\tau)x(\tau) + a_1(\tau)\dot{x}(\tau) + b(\tau)u(\tau),$$

where $u(\tau)$ is a control, which we would like to choose in such a way, that the initial conditions

$$x(0) = x_0, \quad \dot{x}(0) = v_0$$

and the terminal conditions

$$x(1) = x_1, \quad \dot{x}(1) = v_1$$

are satisfied and the functional

$$\int_0^1 \Big(c_0(\tau)x(\tau) + c_1(\tau)\dot{x}(\tau) + d(\tau)u(\tau)\Big)\, d\tau$$

is minimized. In addition, we have a restriction for $u(\tau)$:

$$0 \le u(\tau) \le 1 \quad \text{for all} \quad \tau \in [0,1].$$

This is a typical problem of linear optimal control. The differential equations describes the movement of an "object", thus relating the coordinate $x(\tau)$, the velocity $\dot{x}(\tau)$ and the acceleration $\ddot{x}(\tau)$. The control u is a "force" we may apply. We want to transfer the "object" from the initial position (coordinate, velocity) to the final position (coordinate, velocity), so that the total cost we pay for the coordinate ("gravity"), for the velocity ("friction"), and for the control ("fuel") is minimized.

Our aim is to show that this optimization problem can be considered as a *linear programming problem* of the type:

Find $u \in L^\infty[0,1]$, minimizing $\displaystyle\int_0^1 g(\tau)u(\tau)\, d\tau$

Subject to:

$$\int_0^1 f_1(\tau)u(\tau)\, d\tau = \alpha_1, \quad \int_0^1 f_2(\tau)u(\tau)\, d\tau = \alpha_2, \quad \text{and} \quad 0 \le u(\tau) \le 1,$$

where the functions $f_1, f_2, g \in L^1[0,1]$ and the numbers α_1, α_2 can be computed explicitly.

To see this, let us write the differential equation in the form (28.1.1) by introducing a new variable $v(\tau) = \dot{x}(\tau)$. We get

$$\frac{d}{d\tau}x(\tau) = v(\tau) \quad \text{and} \quad \frac{d}{d\tau}v(\tau) = a_0(\tau)x(\tau) + a_1(\tau)v(\tau) + b(\tau)u(\tau)$$

with the conditions

$$\begin{pmatrix} x(0) \\ v(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x(1) \\ v(1) \end{pmatrix} = \begin{pmatrix} x_1 \\ v_1 \end{pmatrix}.$$

Thus, for vector

$$\mathbf{x}(\tau) = \begin{pmatrix} x(\tau) \\ v(\tau) \end{pmatrix} \quad \text{and} \quad \mathbf{u}(\tau) = \begin{pmatrix} 0 \\ u(\tau) \end{pmatrix}$$

we have

$$A(\tau) = \begin{pmatrix} 0 & 1 \\ a_0(\tau) & a_1(\tau) \end{pmatrix} \quad \text{and} \quad B(\tau) = \begin{pmatrix} 0 & 0 \\ 0 & b(\tau) \end{pmatrix}.$$

Now, the formulas for f_1, f_2 and α_1, α_2 are obtained from (28.1.3). The formula for g can be obtained from (28.1.2).

(28.3) Corollary. *If the problem of optimal linear control is feasible, there exists an optimal solution $u \in L^\infty[0,1]$ such that $u(\tau) \in \{0,1\}$ for almost all $\tau \in [0,1]$. If the optimal solution $u \in L^\infty[0,1]$ is unique, then $u(\tau) \in \{0,1\}$ for almost all $\tau \in [0,1]$.*

Proof. By Proposition 27.3, the set of all feasible controls $u(\tau)$ is weak* compact. By Corollary 26.2, there exists an optimal control u that is an extreme point of the set of all feasible solutions. If u is unique, by Theorem 10.2 (Part 1), u is an extreme point necessarily. By Proposition 27.3, we must have $u(\tau) \in \{0,1\}$ for almost all $\tau \in [0,1]$. □

PROBLEMS.

1°. Consider the equation

$$\ddot{x}(\tau) = a_0(\tau)x(\tau) + a_1(\tau)\dot{x}(\tau) + b(\tau)u(\tau)$$

with the initial conditions

$$x(0) = 0 \quad \text{and} \quad \dot{x}(0) = 0.$$

Show that the solution $x(\tau)$ depends linearly on the control $u(\tau)$: if $x_1(\tau)$ is the solution for $u_1(\tau)$, $x_2(\tau)$ is the solution for $u_2(\tau)$, then $x(\tau) = \alpha_1 x_1(\tau) + \alpha_2 x_2(\tau)$ is the solution for $u(\tau) = \alpha_1 u_1(\tau) + \alpha_2 u_2(\tau)$, where α_1 and α_2 are real numbers.

2°. Let $y(\tau)$ be a solution to the equation

$$\ddot{y}(\tau) = a_0(\tau)y(\tau) + a_1(\tau)\dot{y}(\tau)$$

with some initial conditions

$$y(0) = x_0 \quad \text{and} \quad \dot{y}(0) = v_0$$

and let $z(\tau)$ be a solution of the equation

$$\ddot{z}(\tau) = a_0(\tau)z(\tau) + a_1(\tau)\dot{z}(\tau) + b(\tau)u(\tau)$$

with zero initial conditions. Prove that $x(\tau) = y(\tau) + z(\tau)$ is a solution to the equation

$$\ddot{x}(\tau) = a_0(\tau)x(\tau) + a_1(\tau)\dot{x}(\tau) + b(\tau)u(\tau)$$

with the initial conditions $x(0) = x_0$ and $\dot{x}(0) = v_0$.

Problems 1 and 2 provide some intuition why problems of linear optimal control can be written as linear programs of optimizing $\int_0^1 g(\tau)u(\tau)\,d\tau$ subject to integral constraints $\int_0^1 f_i(\tau)u(\tau)\,d\tau = \alpha_i$ and the "domain" constraint $0 \le u(\tau) \le 1$. First, Problem 2 allows us to reduce (in "nice" cases) the general case to the case of zero initial conditions. Next, if the initial conditions $x(0)$, $\dot{x}(0)$ are zero, the terminal values $x(1)$ and $\dot{x}(1)$ are linear functions of the control u. A "reasonable" linear function (say, weak* continuous) has the form $\int_0^1 f(\tau)u(\tau)\,d\tau$ for some $f \in L^1[0,1]$. Thus fixing the terminal values of the solution x amounts to fixing some integral constraints on u.

29. An Application: The Lyapunov Convexity Theorem

We apply Proposition 27.3 to get a theorem by A.A. Lyapunov (1940) which, in full generality, asserts that the range of a vector non-atomic countably additive measure with values in $\mathbb{R}^n$ is convex. The proof below belongs to J. Lindenstrauss (1966). We adapt it to the special case of a vector-valued measure on $[0,1]$.

(29.1) Theorem. *Let us fix m functions $f_1, \dots, f_m \in L^1[0,1]$. For a Lebesgue measurable subset $A \subset [0,1]$, let $\phi(A) \in \mathbb{R}^m$ be the point*

$$\phi(A) = (\xi_1, \dots, \xi_m), \quad \textit{where} \quad \xi_i = \int_A f_i(\tau)\,d\tau \quad \textit{for} \quad i = 1, \dots, m.$$

Then the set

$$X = \Big\{\phi(A) : \quad A \subset [0,1] \quad \textit{is Lebesgue measurable}\Big\}$$

is a convex compact set in $\mathbb{R}^m$.

Proof. The idea is to prove that X can be represented as the image of a convex compact set under a continuous linear map. Let us extend ϕ to a map $\phi : L^\infty[0,1] \longrightarrow \mathbb{R}^m$,

$$\phi(u) = (\xi_1, \dots, \xi_m), \quad \text{where} \quad \xi_i = \int_0^1 u(\tau)f_i(\tau)\,d\tau \quad \text{for} \quad i = 1, \dots, m.$$

Let

$$B = \Big\{ u \in L^\infty[0,1] : \quad 0 \le u(\tau) \le 1 \quad \text{for almost all} \quad \tau \in [0,1] \Big\}.$$

Then B is convex and weak* compact (see Section 27.2). Since ϕ is linear and continuous, the image $\phi(B) \subset \mathbb{R}^m$ is a compact convex set.

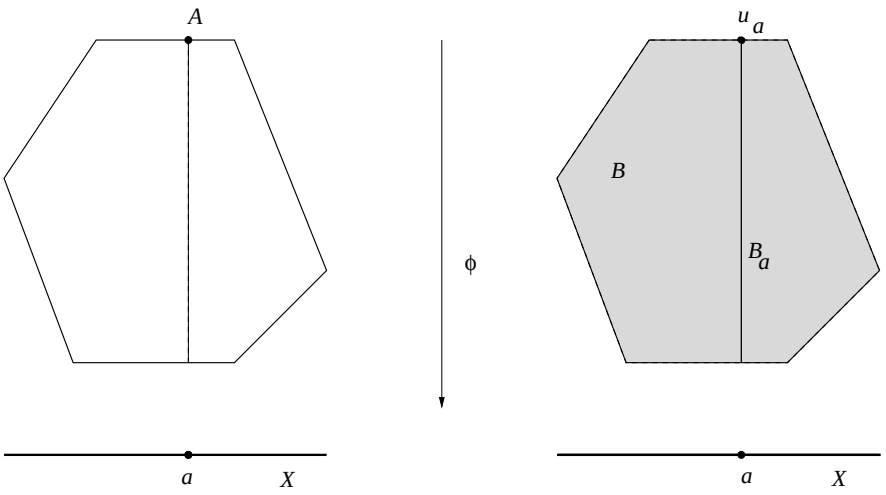


Figure 38

Next, we claim that $\phi(B) = X$. Clearly, $X \subset \phi(B)$.

Let us choose $a = (\alpha_1, \dots, \alpha_m) \in \phi(B)$. Then the set

$$B_a = \Big\{ u \in B : \int_0^1 u(\tau) f_i(\tau)\, d\tau = \alpha_i, \quad \text{for} \quad i = 1, \dots, m \Big\}$$

is non-empty and weak* compact. Therefore, by the Krein-Milman Theorem (Theorem 26.1), there is an extreme point $u_a \in B_a$. By Proposition 27.3, we have $u_a(\tau) \in \{0,1\}$ for almost all τ. Let $A = \big\{ \tau : u_a(\tau) = 1 \big\}$. Then $\phi(A) = a$. Therefore, $\phi(B) \subset X$ and the proof follows. □

Here is an interesting corollary.

(29.2) Corollary. *Let us fix a Lebesgue measurable set $S \subset \mathbb{R}$ and let $F_i(x,\tau)$, $i = 1, \dots, m$ be Lebesgue measurable functions in two real variables $x \in S$ and $\tau \in [0,1]$. Suppose that the functions $F_i(x,\tau)$ are integrable on $S \times [0,1]$. For a Lebesgue measurable function $x(\tau)$ on the interval $[0,1]$ such that $x(\tau) \in S$ for almost all $\tau \in [0,1]$, let*

$$\psi(x) = (\xi_1, \dots, \xi_m), \quad \textit{where} \quad \xi_i = \int_0^1 F_i\big(x(\tau), \tau\big)\, d\tau \quad \textit{for} \quad i = 1, \dots, m.$$

Then the set $X \subset \mathbb{R}^m$*:*

$$X = \Big\{ \psi(x), \quad \textit{where} \quad x(\tau) \quad \textit{is a Lebesgue measurable function on} \quad [0,1]$$

$$\textit{and } x(\tau) \in S \textit{ for almost all } \tau \in [0,1] \Big\}$$

is convex.

Proof. Let us choose two points $a, b \in X$. Hence we have $a = \psi(x)$ and $b = \psi(y)$ for some Lebesgue measurable functions x and y on the interval $[0,1]$, such that $x(\tau), y(\tau) \in S$ for almost all τ. We will construct a convex set $Y \subset \mathbb{R}^m$, such that $Y \subset X$ and $a, b \in Y$. This will prove that X is convex.

To do that, for a Lebesgue measurable subset $A \subset [0,1]$, let us define a function $z_A(\tau)$:

$$z_A(\tau) = \begin{cases} x(\tau) & \text{if } \ \tau \in A \\ y(\tau) & \text{if } \tau \notin A. \end{cases}$$

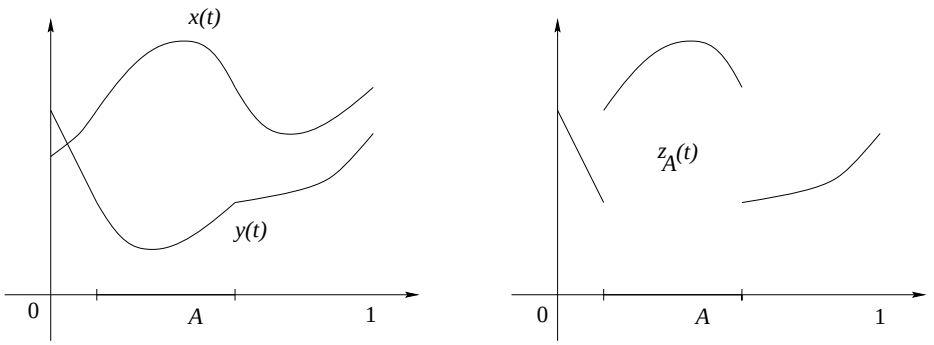


Figure 39

Obviously, z_A is Lebesgue measurable and $z_A(\tau) \in S$ for almost all $\tau \in [0,1]$. Let

$$Y = \Big\{ \psi(z_A) \quad \text{for some Lebesgue measurable} \quad A \subset [0,1] \Big\}.$$

Since $z_{[0,1]}(\tau) = x(\tau)$ and $z_{\emptyset}(\tau) = y(\tau)$, we have $a, b \in Y$. Since $z_A(\tau) \in S$ for almost any τ, we conclude that $Y \subset X$. Let us show that Y is convex. Let us define functions $f_i \in L^1[0,1]$, $i = 1, \ldots, m$ by the formulas:

$$f_i(\tau) = F_i\big(x(\tau), \tau\big) - F_i\big(y(\tau), \tau\big) \quad \text{for} \quad i = 1, \ldots, m.$$

For a Lebesgue measurable set $A \subset [0,1]$, let

$$\phi(A) = (\xi_1, \ldots, \xi_m), \quad \text{where} \quad \xi_i = \int_A f_i(\tau)\, d\tau \quad \text{for} \quad i = 1, \ldots, m.$$

Let

$$c = (\gamma_1, \dots, \gamma_m), \quad \text{where} \quad \gamma_i = \int_0^1 F_i\big(y(\tau), \tau\big)\, d\tau \quad \text{for} \quad i = 1, \dots m.$$

Then $\psi(z_A) = \phi(A) + c$. Therefore, the set Y is a translation (by c) of the set

$$\Big\{\phi(A) : \quad A \subset [0,1] \quad \text{is Lebesgue measurable}\Big\}.$$

The latter set is convex by Theorem 29.1. □

PROBLEMS.

1°. Let us consider the set $X \subset \mathbb{R}^3$ consisting of the points $a = (\alpha_1, \alpha_2, \alpha_3)$, such that

$$\alpha_1 = \int_0^1 \cos\big(\tau x(\tau)\big)\, d\tau, \quad \alpha_2 = \int_0^1 \tau^3 x^2(\tau)\, d\tau, \quad \alpha_3 = \int_0^1 \tau^2 2^{x(\tau)}\, d\tau,$$

where $x(\tau)$ is a Lebesgue measurable function on the interval $[0,1]$, such that $x(\tau) \in (0, e) \cup (\pi, 4)$ for almost all τ. Prove that X is a convex set.

2*. Draw a picture of the set X from Problem 1.

Corollary 29.2 has some interesting applications to problems of optimal control, see A. Megretsky and S. Treil, Power distribution inequalities in optimization and robustness of uncertain systems, *J. Math. Systems Estim. Control*, **3** (1993), no. 3, 301–319.

30. The "Simplex" of Probability Measures

(30.1) The space of continuous functions and its dual. Let $C[0,1]$ be the space of all real-valued continuous functions on the interval $[0,1]$. There is a natural metric on $C[0,1]$: $\operatorname{dist}(f,g) = \max\big\{|f(\tau) - g(\tau)| : \tau \in [0,1]\big\}$, which makes $C[0,1]$ a topological vectors space. The local base at the origin consists of the sets

$$U_\epsilon = \Big\{f \in C[0,1] : |f(\tau)| < \epsilon \quad \text{for all} \quad \tau \in [0,1]\Big\}$$

for $\epsilon > 0$. Let $V[0,1]$ be the space of all continuous linear functionals $\phi : C[0,1] \longrightarrow \mathbb{R}$. The space $V[0,1]$ is often called the *space of signed Borel measures* on [0,1]. The reason for such a name is that every continuous linear functional $\phi : C[0,1] \longrightarrow \mathbb{R}$ can be represented in the form

$$\phi(f) = \int_0^1 f\, d\mu,$$

where μ is a signed Borel measure. For example, μ may be a regular measure, like $\mu = \tau^2\, d\tau$,

$$\phi(f) = \int_0^1 f(\tau)\tau^2\, d\tau$$

or a δ-measure, like $\mu = \delta_{1/2}$:

$$\phi(f) = f(1/2).$$

We make $V[0,1]$ a topological vector space by introducing the weak* topology (see Section 24.9). Thus every continuous linear functional on $V[0,1]$ has the form $\mu \longmapsto \int_0^1 f\, d\mu$ for some fixed $f \in C[0,1]$ and this is the weakest topology with this property.

PROBLEMS.

1. A linear functional $\phi : C[0,1] \longrightarrow \mathbb{R}$ is called *positive*, provided $\phi(f) \geq 0$ for any $f \in C[0,1]$ such that $f(\tau) \geq 0$ for any $\tau \in [0,1]$. Prove that every positive linear functional is continuous.

Hint: Use that $f = \big((|f| + f) - (|f| - f)\big)/2$.

2. Prove that every continuous linear functional $\phi : C[0,1] \longrightarrow \mathbb{R}$ can be represented in the form $\phi = \phi_1 - \phi_2$, where ϕ_1 and ϕ_2 are positive functionals.

Hint: Let $f \in C[0,1]$ be a non-negative function, $f(\tau) \geq 0$ for all $\tau \in [0,1]$. Let $\phi_1(f) = \sup\big\{\phi(g) : 0 \leq g(\tau) \leq f(\tau) \text{ for all } \tau \in [0,1]\big\}$.

3. Let us define the *norm* of a function $f \in C[0,1]$ by $\|f\| = \max\{|f(\tau)| : 0 \leq \tau \leq 1\}$ and the norm of a linear functional $\phi : C[0,1] \longrightarrow \mathbb{R}$ by $\phi(f) = \sup\big\{|\phi(f)| : f \in C[0,1],\ \|f\| = 1\big\}$. Find the norms of the linear functionals:

$$\phi(f) = \int_0^1 f(\tau)\, d\tau, \quad \phi(f) = \int_0^1 \tau f(\tau)\, d\tau, \quad \text{and} \quad \phi_\epsilon(f) = \frac{f(1/2+\epsilon) - f(1/2-\epsilon)}{2\epsilon},$$

where $\epsilon > 0$ is a parameter.

(30.2) Definitions. Let $\Delta \subset V[0,1]$ be the set of all linear functionals $\phi : C[0,1] \longrightarrow \mathbb{R}$ that are positive (that is, $\phi(f) \geq 0$ for all f such that $f(\tau) \geq 0$ for all $\tau \in [0,1]$) and such that $\phi(\mathbf{1}) = 1$, where $\mathbf{1}$ is the function that is identically 1 on $[0,1]$. The set Δ is called the set of all *Borel probability measures on* $[0,1]$. Let us fix a $\tau \in [0,1]$. The linear functional $\delta_\tau : C[0,1] \longrightarrow \mathbb{R}$, where $\delta_\tau(f) = f(\tau)$, is called the *delta measure.*

PROBLEMS.

1°. Prove that Δ is a convex set.

2°. Prove that $\delta_\tau \in \Delta$ for any $\tau \in [0,1]$.

3°. Prove that $\delta_\tau(fg) = \delta_\tau(f)\delta_\tau(g)$ for any $\tau \in [0,1]$ and any two functions $f, g \in C[0,1]$.

4. Let $\phi : C[0,1] \longrightarrow \mathbb{R}$ be a linear functional such that $\phi(fg) = \phi(f)\phi(g)$ for any two functions $f, g \in C[0,1]$. Prove that either $\phi = 0$ or $\phi = \delta_\tau$ for some $\tau \in [0,1]$.

5. Prove that δ_τ is an extreme point of Δ for any $\tau \in [0,1]$.

(30.3) Important fact: The set Δ is compact in the weak* topology (see, for example, W. Rudin, *Functional Analysis*, McGraw-Hill, 1991.

PROBLEM.

1*. Prove that Δ is weak* compact.

Hint: Use the Banach-Alaoglu Theorem, see Problem 4 of Section 24.9.

(30.4) Proposition. *The extreme points of the set Δ of Borel probability measures on the interval $[0,1]$ are the delta-measures δ_τ for $\tau \in [0,1]$.*

Proof. Let us choose a $\tau^* \in [0,1]$. First, we prove that $\delta_* = \delta_{\tau^*}$ is indeed an extreme point of Δ. Clearly, $\delta_* \in \Delta$ (see Problem 2, Section 30.2). Suppose that $\delta_* = (\phi_1 + \phi_2)/2$ for some $\phi_1, \phi_2 \in \Delta$. Let f be any continuous function such that $f(\tau^*) = 1$ and $f(\tau) \leq 1$ for all $\tau \in [0,1]$. Since ϕ_1 and ϕ_2 are positive linear functionals, we have

$$\phi_i(f) = \phi_i\big(\mathbf{1} - (\mathbf{1} - f)\big) = \phi_i(\mathbf{1}) - \phi_i(\mathbf{1} - f) \leq 1.$$

On the other hand $\delta_*(f) = 1$. Therefore, we must have $\phi_1(f) = \phi_2(f) = 1$. By linearity, it follows then that $\phi_1(f) = \phi_2(f) = f(\tau^*)$ for any function $f \in C[0,1]$ such that $f(\tau^*) \geq f(\tau)$ for all $\tau \in [0,1]$. Now, for any $f \in C[0,1]$, we have $f = g_1 - g_2$, where $g_1(\tau) = \min\{f(\tau^*), f(\tau)\}$ and $g_2(\tau) = \min\{0, f(\tau^*) - f(\tau)\}$.

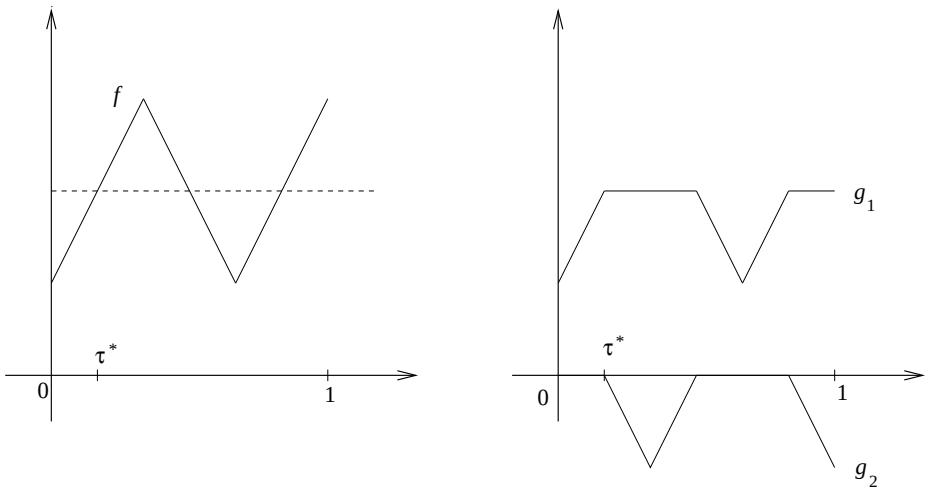


Figure 40

Hence g_1 and g_2 are continuous functions on the interval $[0,1]$ attaining their maximum value at τ^*. Therefore, $\phi_i(g_j) = g_j(\tau^*)$ for $i, j = 1, 2$. By linearity, we conclude that $\phi_i(f) = \phi_i(g_1) - \phi_i(g_2) = f(\tau^*)$. Therefore, $\phi_1 = \phi_2 = \delta_*$, so δ_* is an extreme point.

Suppose that ϕ is an extreme point of Δ. First, we establish that $\phi(fg) = \phi(f)\phi(g)$ for any two functions $f, g \in C[0,1]$ and then we deduce that $\phi = \delta_\tau$ for some $\tau \in [0,1]$. Let us fix a function $h \in C[0,1]$. Then $\psi : C[0,1] \longrightarrow \mathbb{R}$ defined by $\psi(f) = \phi(hf)$ is a continuous linear functional on $C[0,1]$. Let us choose h such that $0 < h(\tau) < 1$ for all $\tau \in [0,1]$. Since ϕ is positive, we have $1 > \phi(h) > 0$. We define $\psi_1, \psi_2 : C[0,1] \longrightarrow \mathbb{R}$ by

$$\psi_1(f) = \frac{\phi(hf)}{\phi(h)} \quad \text{and} \quad \psi_2(f) = \frac{\phi\big((\mathbf{1} - h)f\big)}{\phi(\mathbf{1} - h)}.$$

It is easy to see that $\psi_i \in \Delta$. We can write ϕ as a convex combination

$$\phi = \phi(h)\psi_1 + \phi(\mathbf{1} - h)\psi_2.$$

Since ϕ is an extreme point of Δ, we must have $\psi_1 = \psi_2 = \phi$. In particular, $\phi(fh) = \phi(h)\phi(f)$ for any $f \in C[0,1]$ and any h such that $0 < h(\tau) < 1$ for each $\tau \in [0,1]$. By linearity, it follows that $\phi(fg) = \phi(f)\phi(g)$ for any two $f, g \in C[0,1]$.

Let $H = \big\{f \in C[0,1] : \phi(f) = 0\big\}$ be the kernel of ϕ. Hence H is a (closed) hyperplane in $C[0,1]$. We observe that for every $f \in H$, there is a $\tau \in [0,1]$ such that $f(\tau) = 0$. Indeed, if $f \in H$ is a function which is nowhere 0, then $1/f$ is a continuous function and we would have had $1 = \phi(\mathbf{1}) = \phi(f \cdot (1/f)) = \phi(f)\phi(1/f) = 0$, which is a contradiction. Next, we observe that any set of finitely many functions $f_1, \ldots, f_m \in H$ has a common zero τ^*: $f_1(\tau^*) = \ldots = f_m(\tau^*) = 0$. Otherwise, the function $f = f_1^2 + \ldots + f_m^2$ is everywhere positive and $f \in H$ since $\phi(f) = \phi(f_1^2) + \ldots + \phi(f_m^2) = \phi^2(f_1) + \ldots + \phi^2(f_m) = 0$. But we already proved that every function $f \in H$ must have a zero. Finally, we conclude that there is a point $\tau^* \in [0,1]$ such that $f(\tau^*) = 0$ for any $f \in H$. Indeed, for any $f \in C[0,1]$, the set $X_f = \big\{\tau \in [0,1] : f(\tau) = 0\big\}$ is a closed set and any finite intersection $X_{f_1} \cap \ldots \cap X_{f_m}$ for $f_1, \ldots, f_m \in H$ is non-empty. Since $[0,1]$ is a compact interval, the intersection $\bigcap_{f \in H} X_f$ is non-empty. Now we see that H is a subset of the kernel of δ_{τ^*}, which implies that $\phi = \alpha\delta_{\tau^*}$ for some $\alpha \in \mathbb{R}$. Hence $\phi = \delta_{\tau^*}$ and the proof follows. $\square$

Remark. The set Δ can be considered as an (uncountably) infinite-dimensional analogue of the standard d-dimensional simplex $\Delta_d \subset \mathbb{R}^{d+1}$:

$$\Delta_d = \Big\{(\gamma_0, \ldots, \gamma_d) : \gamma_0 + \ldots + \gamma_d = 1 \quad \text{and} \quad \gamma_i \geq 0 \quad \text{for} \quad i = 0, \ldots, d\Big\}.$$

The analogy is based on the following observation: the "discretization" of the space $C[0,1]$ is the space $C[\tau_0, \ldots, \tau_d]$ of all continuous functions on the discrete set of $d+1$ points $0 \leq \tau_0 < \tau_1 < \ldots < \tau_d \leq 1$. The space $C[\tau_0, \ldots, \tau_d]$ can be identified with $\mathbb{R}^{d+1}$, the space of all $(d+1)$-tuples $(\xi_0, \ldots, \xi_d)$. The "discretization" of $V[0,1]$ is the space of all continuous linear functionals $\phi : C[\tau_0, \ldots, \tau_d] \longrightarrow \mathbb{R}$, which is

identified with the space $\mathbb{R}^{d+1}$ of all $(d+1)$-tuples $(\gamma_0, \ldots, \gamma_d)$: $\phi(\xi_0, \ldots, \xi_d) = \gamma_0\xi_0 + \ldots + \gamma_d\xi_d$. Then the "discretization" of $\Delta \subset V[0,1]$ is the simplex Δ_d.

Notation. For $\mu \in V[0,1]$ and $f \in C[0,1]$, instead of $\mu(f)$ we often write $\int_0^1 f\, d\mu$.

PROBLEM.

1. Let $O \subset V[0,1]$ be the set of functionals ϕ such that $|\phi(f)| \leq 1$ for any $f \in C[0,1]$ with the property that $|f(\tau)| \leq 1$ for each $\tau \in [0,1]$. Prove that the extreme points of O are δ_τ and $-\delta_\tau$ for $\tau \in [0,1]$.

31. Extreme Points of an Intersection. Applications

We need a simple and useful result which describes the extreme points of the intersection of a convex set with an affine subspace. First, we describe the intersection with a hyperplane.

(31.1) Lemma. *Let V be a vector space and let $K \subset V$ be a convex set, such that for every straight line $L \subset V$, the intersection $K \cap L$ is a closed bounded interval, possibly empty or a point. Let $H \subset V$ be an affine hyperplane. Then every extreme point of $K \cap H$ can be expressed as a convex combination of at most two extreme points of K.*

Proof. Let u be an extreme point of $K \cap H$. If u is an extreme point of K, the proof follows. Otherwise, there are two distinct points $u_+, u_- \in K$, such that $u = (u_+ + u_-)/2$. Let L be the straight line passing through u_+ and u_-. The intersection $L \cap K = [u_1, u_2]$ is a closed interval, containing u in its interior. We claim that u_1 and u_2 are extreme points of K. Suppose, for example, that u_1 is not an extreme point of K. Then there are two distinct points $v_+, v_- \in K$, such that $u_1 = (v_+ + v_-)/2$. Clearly, $v_+ \notin L$ and $v_- \notin L$. Consider the 2-dimensional plane A passing through the points v_+, v_-, and u_2 and the triangle $\Delta = \mathrm{conv}(v_+, v_-, u_2)$ in the plane A.

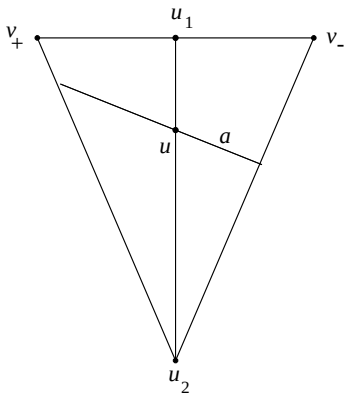


Figure 41

The point u is an interior point of the triangle. Since H is a hyperplane, the intersection $a = A \cap H$ is a straight line, passing through u. Since u is an interior point of Δ, the intersection $\Delta \cap H$ is an interval, containing u as its interior point. Since K is convex, $\Delta \cap H \subset K \cap H$, so the intersection $K \cap H$ contains an interval, containing u as its interior point, which contradicts the assumption that u is an extreme point of $K \cap H$.

The contradiction shows that u_1 and u_2 are extreme points of $K \cap H$. Since u is a convex combination of u_1 and u_2, the proof follows. □

(31.2) Theorem. *Let V be a vector space and let $K \subset V$ be a convex set, such that for every straight line $L \subset V$, the intersection $K \cap L$ is a closed bounded interval, possibly empty or a point. Let $H_1, \ldots, H_m \subset V$ be a finite set of affine hyperplanes in V. Then, every extreme point of the intersection $K \cap H_1 \cap \ldots \cap H_m$ can be expressed as a convex combination of at most $m+1$ extreme points of K. Equivalently, if $A \subset V$ is an affine subspace such that* $\operatorname{codim} A = m$, *then every extreme point of the intersection $K \cap A$ can be expressed as a convex combination of at most $m+1$ extreme points of K.*

Proof. First, we prove that every extreme point of the intersection $K \cap H_1 \cap \ldots \cap H_m$ is a convex combination of extreme points of K. Let us define $K_0 = K$, $K_1 = K \cap H_1$, $K_i = K_{i-1} \cap H_i$, $i = 1, \ldots, m$. Since the intersection of a straight line L with any affine subspace in V is either L itself, or a point, or empty, we can apply Lemma 31.1 to K_i. Hence we conclude that every extreme point of K_i is a convex combination of at most 2 extreme points of K_{i-1}. Therefore, every extreme point of $K_m = K \cap H_1 \cap \ldots \cap H_m$ is a convex combination of at most 2^m extreme points of $K_0 = K$.

Let u be an extreme point of $K \cap H_1 \cap \ldots \cap H_m$. Let us choose a representation

$$u = \alpha_1 u_1 + \ldots + \alpha_n u_n, \quad \text{where} \quad \alpha_1 + \ldots + \alpha_n = 1, \quad \alpha_i \geq 0 \quad \text{and} \quad u_i \in \operatorname{ex}(K)$$

for $i = 1, \ldots, n$ with the smallest possible n (we know that we can choose $n \leq 2^m$). Clearly, $\alpha_i > 0$ for all $i = 1, \ldots, n$. Furthermore, the points $u_1, \ldots, u_n$ are affinely independent, since otherwise we could have reduced n as in the proof of Carathéodory's Theorem (see Theorem 2.3). Therefore, u is an interior point of the $(n-1)$-dimensional simplex $\Delta = \operatorname{conv}(u_1, \ldots, u_n) \subset K$ (cf. Problem 1, Section 9.3). If $n > m+1$, then the intersection $\Delta \cap H_1 \ldots \cap H_m$ contains an interval containing u as its interior point, which contradicts the assumption that u is an extreme point of $K \cap H_1 \cap \ldots \cap H_m$. Hence $n \leq m+1$ and the proof follows. □

PROBLEMS.

1. Show by example that, in general, the constant $m+1$ in Theorem 21.2 can not be reduced.

2°. Let $K = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ be a ball. Prove that every extreme point of $K \cap H_1 \cap \ldots \cap H_m$ is an extreme point of K.

3. Let $\mathbb{R}^d : d = n(n+1)/2$ be the space of $n \times n$ symmetric matrices and let $K = \{X \succeq 0 \text{ and } \operatorname{tr}(X) = 1\}$. Prove that every extreme point of $K \cap H_1 \cap \ldots \cap H_m$

is a convex combination of not more than $\lfloor(\sqrt{8m+1}-1)/2\rfloor+1$ extreme points of K.

Hint: Cf. Proposition 20.1.

4. Let $\mathbb{R}^d : d = n^2$ be the space of all $n \times n$ matrices and let $K \subset \mathbb{R}^d$ be the Birkhoff Polytope (the polytope of doubly stochastic matrices, see Section 12). Let $S \subset \mathbb{R}^d$ be the subspace of all $n \times n$ symmetric matrices. Prove that every extreme point of $K \cap S$ can be represented as a convex combination of not more than two extreme points of K.

A generalization of Problem 4:

5. Let G be a finite group of linear transformations of $\mathbb{R}^d$. Let $K \subset \mathbb{R}^d$ be a convex compact set, which is G-invariant: $g(x) \in K$ for all $g \in G$ and all $x \in K$. Let $L = \{x \in \mathbb{R}^d : g(x) = x \text{ for all } g \in G\}$ be the subspace of G-invariant vectors (check, that L is indeed a subspace). Prove that every extreme point of $K \cap L$ is a convex combination of not more than $|G|$ (the cardinality of G) extreme points of K.

Theorem 31.2 looks intuitively obvious in small dimensions.

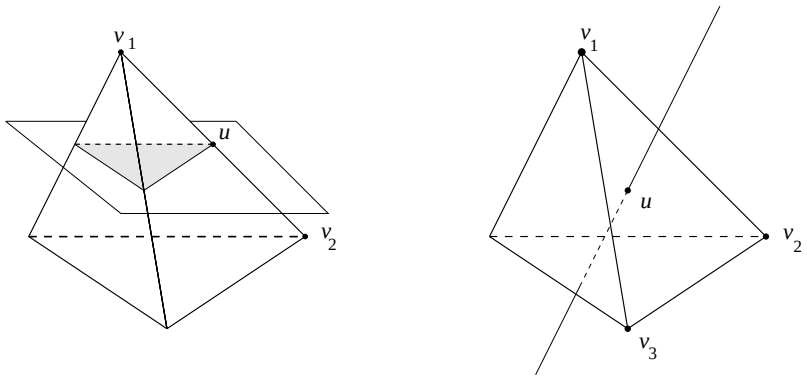


Figure 42

It has some interesting infinite-dimensional applications.

(31.3) Application: extreme points of a set of probability measures. Let μ be an (unknown) Borel probability measure on the interval $[0,1]$ (see Section 30) and let $f_1, \dots, f_m \in C[0,1]$ be continuous functions on the interval $[0,1]$. Suppose that the measure μ is not known, but we know the expectations

$$\int_0^1 f_i \, d\mu = \alpha_i, \quad i = 1, \dots, m$$

and the functions $f_1, \dots, f_m$. We want to estimate the expectation

$$\alpha_0 = \int_0^1 f_0 \, d\mu$$

of yet another known function $f_0 \in C[0,1]$. Quite often, we have $f_i(\tau) = \tau^i$, $i = 1, \ldots, m$, so α_i are *moments* of an unknown probability distribution, and f_0 is some other function of interest. Theorem 31.2 and Proposition 30.4 lead to the following useful *discretization principle*.

(31.4) Proposition. *Let us fix continuous functions $f_0, f_1, \ldots, f_m$ on the interval $[0,1]$ and real numbers $\alpha_1, \ldots, \alpha_m$. Suppose that the set B of Borel probability measures μ on $[0,1]$ satisfying the equations*

$$\int_0^1 f_i \, d\mu = \alpha_i \quad \textit{for} \quad i = 1, \ldots, m$$

is non-empty. Then there exist measures $\mu^+, \mu^- \in B$ such that

(1) *The measures μ^+ and μ^- are convex combinations of at most $m+1$ delta-measures:*

$$\mu^+ = \sum_{i=1}^{m+1} \lambda_i^+ \delta_{\tau_i^+}, \quad \mu^- = \sum_{i=1}^{m+1} \lambda_i^- \delta_{\tau_i^-},$$

where

$$\lambda_i^+, \lambda_i^- \geq 0 \quad \textit{for} \quad i = 1, \ldots, m+1, \qquad \sum_{i=1}^{m+1} \lambda_i^+ = \sum_{i=1}^{m+1} \lambda_i^- = 1$$

and $\tau_i^+, \tau_i^- \in [0,1]$ for $i = 1, \ldots, m+1$;

(2) *The set of values*

$$\alpha_0 = \int_0^1 f_0 \, d\mu \quad \textit{for} \quad \mu \in B$$

is the interval $[\alpha^-, \alpha^+]$, where

$$\alpha^- = \int_0^1 f_0 \, d\mu^- \quad \textit{and} \quad \alpha^+ = \int_0^1 f_0 \, d\mu^+.$$

Proof. The set B can be represented as the intersection $B = \Delta \cap H_1 \cap \ldots \cap H_m$, where Δ is the simplex of Borel probability measures (see Section 30) and the affine hyperplanes H_i are defined by the equations:

$$H_i = \{\mu \in V[0,1] : \int_0^1 f_i \, d\mu_i = \alpha_i\}.$$

Clearly, B is convex. Since Δ is weak* compact (Fact 30.3) and H_i are closed in the weak* topology, the set B is compact. The function

$$\alpha_0 : \mu \longmapsto \int_0^1 f \, d\mu$$

is weak* continuous and hence by Corollary 26.2 there is an extreme point μ^- of B where α_0 attains its minimum on B and there is an extreme point μ^+ of B where α_0 attains its maximum on B. By Theorem 31.2, μ^- and μ^+ can be represented as convex combinations of some $m+1$ extreme points of Δ. Proposition 30.4 implies that the extreme points of Δ are the delta-measures. □

PROBLEMS.

1. Let us fix an $\alpha \in [0,1]$. Find a probability measure μ on $[0,1]$ such that $\int_0^1 \tau \, d\mu = \alpha$ and $\int_0^1 (\tau - a)^2 \, d\mu$ is maximized.

2. Let μ be a probability measure on $[0,1]$. Prove that for the *variance*

$$D(\mu) = \int_0^1 \tau^2 \, d\mu - \Big(\int_0^1 \tau \, d\mu\Big)^2$$

one has $0 \le D(\mu) \le 1/4$.

Hint: Use Problem 1.

3. Let us consider the set S_{d+1} of all $(d+1)$-tuples $(\xi_0, \dots, \xi_d)$, where $\xi_i = \int_0^1 \tau^i \, d\mu$ for some $\mu \in \Delta$ and $i = 0, \dots, d$. Prove that S_{d+1} is the section of the moment cone M_{d+1} (see Section 16) by the hyperplane $\xi_0 = 1$.

32. Polarity in Euclidean Space

(32.1) Definition. Let $A \subset \mathbb{R}^d$ be a set. The set

$$A^\circ = \Big\{ c \in \mathbb{R}^d : \langle c, x \rangle \le 1 \quad \text{for all} \quad x \in A \Big\}$$

is called the *polar* of A.

PROBLEMS.

1°. Prove that A° is a convex closed set containing the origin.

2°. Prove that $\big(\mathbb{R}^d\big)^\circ = \{0\}$ and that $\{0\}^\circ = \mathbb{R}^d$.

3°. Prove that if $A \subset B$ then $B^\circ \subset A^\circ$.

4°. Prove that $\Big(\bigcup_{i \in I} A_i\Big)^\circ = \bigcap_{i \in I} A_i^\circ$.

5°. Let $A \subset \mathbb{R}^d$ be a set and let $\alpha > 0$ be a number. Let $\alpha A = \big\{\alpha x : x \in A\big\}$. Prove that $(\alpha A)^\circ = \alpha^{-1} A^\circ$.

6. Let $L \subset \mathbb{R}^d$ be a linear subspace. Prove that L° is the orthogonal complement of L.

7°. Let $A = \operatorname{conv}\big(v_1, \dots, v_m\big)$ be a polytope. Prove that $A^\circ = \big\{x : \langle v_i, x \rangle \le 1, i = 1, \dots, m\big\}$.

8°. Prove that $A \subset (A^\circ)^\circ$.

(32.2) The Bipolar Theorem. *Let $A \subset \mathbb{R}^d$ be a convex closed set containing the origin. Then $(A^\circ)^\circ = A$.*

Proof. From Problem 8°, Section 32.1, we conclude that $A \subset (A^\circ)^\circ$. It remains to show, therefore, that $(A^\circ)^\circ \subset A$. Suppose there is a point u such that $u \in (A^\circ)^\circ$ and yet $u \notin A$. Since A is closed and convex, by Theorem 23.3 there exists an affine hyperplane strictly separating u form A. In other words, there exists a vector $c \neq 0$ and a number α such that $\langle c, x\rangle < \alpha$ for all $x \in A$ and $\langle c, u\rangle > \alpha$. Since $0 \in A$, we conclude that $\alpha > 0$. Let us consider $b = \alpha^{-1}c$. We have $\langle b, x\rangle < 1$ for all $x \in A$ and hence $b \in (A^\circ)$. However, $\langle b, u\rangle > 1$, which contradicts the assumption that $u \in (A^\circ)^\circ$. Therefore $(A^\circ)^\circ \subset A$ and the proof follows. □

PROBLEMS.

1°. Prove that $((A^\circ)^\circ)^\circ = A^\circ$ for every non-empty set $A \subset \mathbb{R}^d$.

2. Let $A \subset \mathbb{R}^d$ be a set. Prove that $(A^\circ)^\circ$ is the closure of $\operatorname{conv}\bigl(A, 0\bigr)$.

3. Let $A = \bigl\{x \in \mathbb{R}^d : \langle c_i, x\rangle \leq 1 \text{ for } i = 1, \ldots, m\bigr\}$. Prove that $A^\circ = \operatorname{conv}(0, c_1, \ldots, c_m)$.

4. Let $A \subset \mathbb{R}^d$ be a set such that $A^\circ = A$. Prove that $A = \bigl\{x : \|x\| \leq 1\bigr\}$ is the unit ball.

5. Let $A = \bigl\{x \in \mathbb{R}^d : -1 \leq \xi_i \leq 1 \text{ for } i = 1, \ldots, d\bigr\}$ and let $B = \bigl\{x \in \mathbb{R}^d : |\xi_1| + \ldots + |\xi_d| \leq 1\bigr\}$. Prove that $A^\circ = B$ and $B^\circ = A$.

6. Let us fix $p, q > 0$ such that $1/p + 1/q = 1$. Let $A = \bigl\{x \in \mathbb{R}^d : \sum_{i=1}^d |\xi_i|^p \leq 1\bigr\}$ and let $B = \bigl\{x \in \mathbb{R}^d : \sum_{i=1}^d |\xi_i|^q \leq 1\bigr\}$. Prove that $A^\circ = B$ and that $B^\circ = A$.

7°. Let $B_\lambda = \bigl\{x \in \mathbb{R}^d : \|x\| \leq \lambda\bigr\}$ be the ball of the radius λ. Prove that $(B_\lambda)^\circ = B_{1/\lambda}$.

Now we are ready to prove the second part of the Weyl-Minkowski Theorem (the first part is Corollary 11.3).

(32.3) Corollary. *A polytope is a polyhedron.*

Proof. Let $P = \operatorname{conv}\bigl(v_1, \ldots, v_m\bigr)$ be a polytope in $\mathbb{R}^d$. Without loss of generality we may assume that $\operatorname{int} P \neq \emptyset$ (otherwise, we consider the smallest affine subspace containing P, see Theorem 9.4) and that $0 \in \operatorname{int} P$ (otherwise, we shift P). In other words, $B_\epsilon \subset P$ for some $\epsilon > 0$, where B_λ is the ball of the radius λ centered at the origin. By Problem 3, Section 32.1 and Problem 7, Section 32.2, we have $P^\circ \subset B_\epsilon^\circ = B_{1/\epsilon}$, so P° is bounded. Furthermore,

$$P^\circ = \Bigl\{c : \langle c, v_i\rangle \leq 1 \quad \text{for} \quad i = 1, \ldots, m\Bigr\},$$

so P° is a polyhedron (cf. Problem 7, Section 32.1). Hence P° is a bounded polyhedron, and, therefore, by Corollary 11.3 is a polytope:

$$P^\circ = \operatorname{conv}\bigl(u_1, \ldots, u_n\bigr)$$

for some $u_i \in \mathbb{R}^d$. Applying the Bipolar Theorem (Theorem 32.2), we conclude that

$$P = (P^\circ)^\circ = \left\{ x \in \mathbb{R}^d : \langle x, u_i \rangle \leq 1 \quad \text{for} \quad i = 1, \ldots, n \right\}$$

is a polyhedron. □

PROBLEM.

1. Prove that the polar of a polyhedron is a polyhedron.

Some interesting dualities:

2. Let us define the *projective plane* $\mathbb{RP}^2$ as follows: the points of $\mathbb{RP}^2$ are the straight lines in $\mathbb{R}^3$, passing through the origin. The straight lines in $\mathbb{RP}^2$ are the planes in $\mathbb{R}^3$ passing through the origin. As usual, a line in $\mathbb{RP}^2$ (a plane in $\mathbb{R}^3$) consists of points of $\mathbb{RP}^2$ (lines in $\mathbb{R}^3$).

Prove that for every two distinct points A and B of $\mathbb{RP}^2$, there is a unique straight line $c \subset \mathbb{RP}^2$ that contains A and B.

Prove that every two distinct lines $a, b \in \mathbb{RP}^2$ intersect at a unique point C.

There is the *polarity correspondence* between points and lines of $\mathbb{RP}^2$: a point $A \in \mathbb{RP}^2$ corresponds to a line $a \subset \mathbb{RP}^2$ if and only if A, considered as a line in $\mathbb{R}^3$, is orthogonal to a, considered as a plane in $\mathbb{R}^3$. Prove that a point C is the intersection of two distinct lines a and b if and only if the distinct points A and B lie on the straight line c.

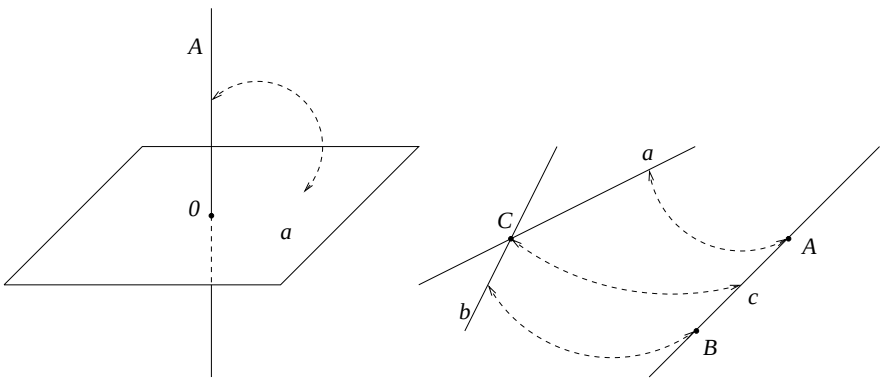


Figure 43

The Euclidean plane $\mathbb{R}^2$ can be embedded in $\mathbb{RP}^2$ in the following way. Let us identify $\mathbb{R}^2$ with an affine plane in $\mathbb{R}^3$, not passing through the origin. A point $x \in \mathbb{R}^2$ is identified with the straight line $X \subset \mathbb{R}^3$, passing through x and the

origin. Hence, x is identified with a point of $\mathbb{RP}^2$.

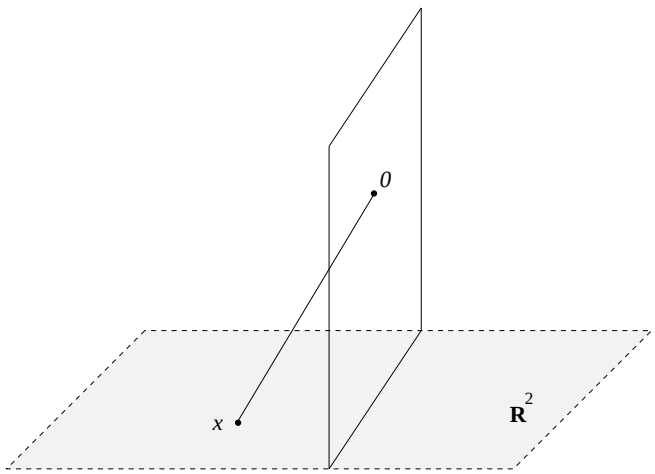


Figure 44

Prove that the straight lines in $\mathbb{R}^2$ are identified with straight lines in $\mathbb{RP}^2$. Describe the points of $\mathbb{RP}^2$ that are not identified with any point of $\mathbb{R}^2$ and the straight lines in $\mathbb{RP}^2$ that are not identified with any straight line of $\mathbb{R}^2$.

3. A polytope $P \subset \mathbb{R}^d$ is called *self-dual*, if it contains the origin as its interior point and P° can be obtained from P by a non-degenerate linear transformation. Suppose that $P = \operatorname{conv}(v_1, \ldots, v_{d+1})$, where $v_1, \ldots, v_{d+1}$ are affinely independent and $0 = v_1 + \ldots + v_{d+1}$. Prove that P is self-dual.

4. Prove that for $d > 2$, the cube $I_d = \big\{(\xi_1, \ldots, \xi_d) : -1 \le \xi_i \le 1, i = 1, \ldots, d\big\}$ is not self-dual. Prove that I_1 and I_2 are self-dual.

5. Prove that a regular polygon in $\mathbb{R}^2$, containing the origin as its center, is self-dual.

6. Let e_1, e_2, e_3, e_4 be the standard orthonormal basis of $\mathbb{R}^4$. The polytope

$$P = \operatorname{conv}\big(e_i + e_j, -e_i - e_j, e_i - e_j : \quad 1 \le i \ne j \le 4\big)$$

is called the *24-cell*. Prove that P is self-dual.

Polytope P is a regular 4-dimensional polytope having 24 vertices and 24 facets.

Recall (Section 15.1) that a convex set $K \subset \mathbb{R}^d$ is called a convex *cone* if $0 \in K$ and $\lambda x \in K$ for every $x \in K$ and $\lambda \ge 0$.

(32.4) Lemma. *Let $K \subset \mathbb{R}^d$ be a convex cone. Then*

$$K^\circ = \Big\{x \in \mathbb{R}^d : \langle x, y\rangle \le 0 \quad \textit{for every} \quad y \in K\Big\}.$$

Proof. By Definition 32.1, K° consists of all the points x in $\mathbb{R}^d$ such that $\langle x, y\rangle \le 1$ for all $y \in K$. Suppose that for some $x \in K^\circ$ and some $y \in K$ we have $\langle x, y\rangle > 0$. Then, for a sufficiently large $\lambda > 0$, one has $\langle x, \lambda y\rangle > 1$. Since K is a cone, $\lambda y \in K$, which contradicts the definition of K°. □

PROBLEMS.

1°. Prove that the polar of a cone is a cone.

2°. Let $K_1, K_2 \subset \mathbb{R}^d$ be cones. Prove that $(K_1 + K_2)^\circ = K_1^\circ \cap K_2^\circ$.

3. Let $K_1, K_2 \subset \mathbb{R}^d$ be convex closed cones. Prove that $(K_1 \cap K_2)^\circ$ is the closure of $K_1^\circ + K_2^\circ$.

4. Prove that a convex cone $K \subset \mathbb{R}^d$ is the conic hull of a finite set if and only if K can be represented in the form $K = \{x \in \mathbb{R}^d : \langle c_i, x\rangle \le 0, i \in I\}$ for finite, possibly empty, set I of vectors $c_i \in \mathbb{R}^d$. Such cones are called *polyhedral*.

5°. Prove that a non-empty face of a cone is a cone.

It turns out that polarity can be extended to a valuation (see Definition 6.3), when restricted to cones. Recall (see Definition 6.3) that $\mathcal{C}(\mathbb{R}^d)$ is the vector space (algebra) spanned by the indicator functions of closed convex sets in $\mathbb{R}^d$. Since an intersection of cones is a cone, the subspace spanned by the indicator functions of closed cones is a subalgebra of $\mathcal{C}(\mathbb{R}^d)$.

(32.5) Theorem. *Let $\mathcal{C}o(\mathbb{R}^d)$ be the subalgebra of the algebra $\mathcal{C}(\mathbb{R}^d)$ of closed convex sets spanned by the indicator functions $[K]$ of convex closed cones $K \subset \mathbb{R}^d$. There exists a linear transformation $\mathcal{D} : \mathcal{C}o(\mathbb{R}^d) \longrightarrow \mathcal{C}o(\mathbb{R}^d)$ such that $\mathcal{D}([K]) = [K^\circ]$ for any closed convex cone $K \subset \mathbb{R}^d$.*

Proof. Let us consider a function $F : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}$,

$$F(x, y) = \begin{cases} 1 & \text{if } \langle x, y\rangle = 1 \\ 0 & \text{otherwise.} \end{cases}$$

For a function $g \in \mathcal{C}(\mathbb{R}^d)$ and a fixed $y \in \mathbb{R}^d$, let us consider the function $h_y(x) = g(x)F(x, y)$ as a function of $x \in \mathbb{R}^d$. We claim that $h_y \in \mathcal{C}(\mathbb{R}^d)$. By linearity, it suffices to check this in the case when $g = [A]$ is the indicator function of a closed convex set $A \subset \mathbb{R}^d$. In this case, $h_y = [A \cap H_y]$, where

$$H_y = \{x \in \mathbb{R}^d : \langle x, y\rangle = 1\}$$

is a hyperplane. Since $A \cap H_y$ is a closed convex set, $h_y \in \mathcal{C}(\mathbb{R}^d)$.

Since $h_y(x) \in \mathcal{C}(\mathbb{R}^d)$, we can apply the Euler characteristic μ (see Section 6) to h_y. Now we are ready to define the operator $\mathcal{D}$. For $g \in \mathcal{C}o(\mathbb{R}^d)$, let $f = \mathcal{D}(g)$ be defined by the formula

$$f(y) = \mu(g) - \mu(h_y) = \mu(g) - \mu\big(g(x)F(x, y)\big).$$

Using Theorem 6.4, we conclude that $\mathcal{D}(\alpha_1 g_1 + \alpha_2 g) = \alpha_1 \mathcal{D}(g_1) + \alpha_2 \mathcal{D}(g_2)$, so $\mathcal{D}$ is linear. It remains to check that $\mathcal{D}([K]) = [K^\circ]$ for any closed convex cone K. Indeed, if $g = [K]$ then $h_y = g(x)F(x,y) = [K \cap H_y]$, so for $f = \mathcal{D}(g)$ we have

$$f(y) = \begin{cases} 0 & \text{if } K \cap H_y \neq \emptyset \\ 1 & \text{if } K \cap H_y = \emptyset. \end{cases}$$

If $y \in K^\circ$ then by Lemma 32.4 $\langle x, y \rangle \leq 0$ for all $x \in K$, so $K \cap H_y = \emptyset$ and $f(y) = 1$. If $y \notin K^\circ$ then by Lemma 32.4 there exists an $x \in K$ such that $\langle x, y \rangle > 0$. Hence for some $\lambda > 0$ we have $\langle \lambda x, y \rangle = 1$. Since K is a cone, $\lambda x \in K$ and hence $K \cap H_y \neq \emptyset$. Therefore, $f(y) = 0$ and the proof follows. $\square$

PROBLEMS.

1°. Prove that $\mathcal{D}^2 = \mathrm{id}$, where id is the identical transformation of $\mathcal{C}o(\mathbb{R}^d)$.

2. Find the eigenvalues of $\mathcal{D}$ as an operator $\mathcal{D} : \mathcal{C}o(\mathbb{R}^d) \longrightarrow \mathcal{C}o(\mathbb{R}^d)$.

3. Consider the subalgebra $\mathcal{C}op(\mathbb{R}^d)$ generated by the indicator functions $[K]$ of polyhedral cones K (see Problem 4 of Section 32.4). Prove that there exists a bilinear operation $\star$ such that $[K_1] \star [K_2] = [K_1 + K_2]$ for any two polyhedral cones K_1 and K_2 (cf. Problem 1 of Section 7.2). Prove that the operator $\mathcal{D}$ maps the space $\mathcal{C}op(\mathbb{R}^d)$ onto itself. Furthermore, prove that $\mathcal{D}(fg) = \mathcal{D}(f) \star \mathcal{D}(g)$. Thus, polarity plays the role of the Fourier transform with the Euler characteristic as a measure.

Theorem 32.5 implies that if the indicator functions of some convex closed cones satisfy a linear relation, then the indicator functions of their polars satisfy the same

linear relation.

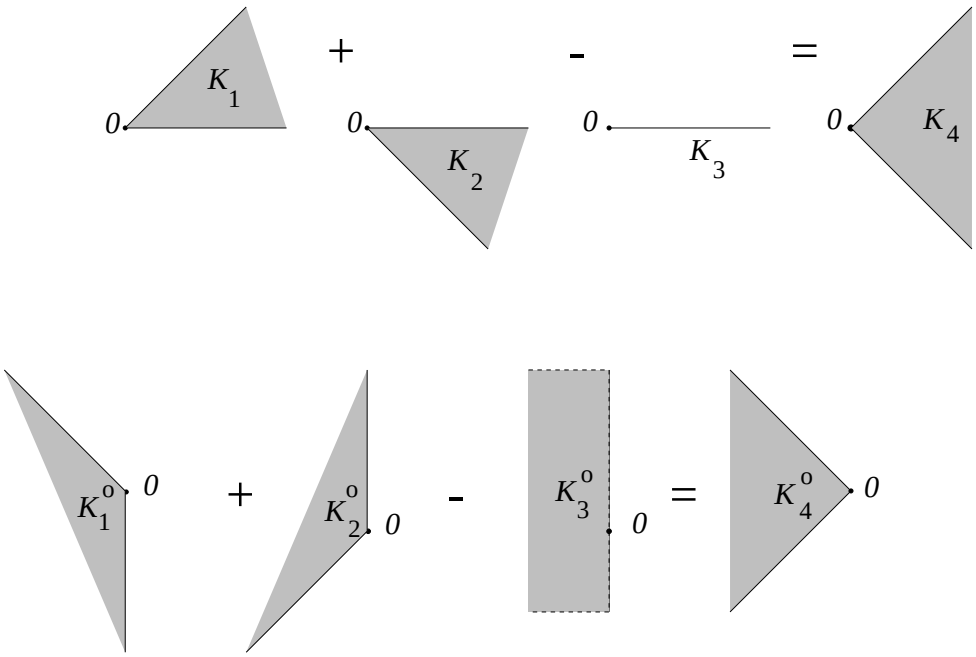


Figure 45

(32.6) Corollary. *Let $K_i : i = 1, \dots, m$ be closed convex cones in $\mathbb{R}^d$ and let $\alpha_i : i = 1, \dots, m$ be real numbers such that*

$$\sum_{i=1}^{m} \alpha_i [K_i] = 0.$$

Then

$$\sum_{i=1}^{m} \alpha_i [K_i^\circ] = 0.$$

Proof. Applying the operator $\mathcal{D}$ of Theorem 32.5 to both sides of the identity $\sum_{i=1}^{m} \alpha_i [K_i] = 0$. □

33. An Application: Recognizing Points in the Moment Cone

As an application of the Bipolar Theorem (Theorem 32.2), we show how to decide whether a given point belongs to the moment cone (see Section 16).

Let us consider the space $\mathbb{R}^{d+1}$ of all $d+1$-tuples $x = (\xi_0, \dots, \xi_d)$. Let us fix the interval $[0, 1] \subset \mathbb{R}$ (the case of a general interval $[\alpha, \beta] \subset \mathbb{R}$ is treated similarly). Let

$$g(\tau) = (1, \tau, \tau^2, \dots, \tau^d) : \quad 0 \leq \tau \leq 1$$

be the moment curve and let

$$M = \operatorname{co}\big(g(\tau) : 0 \le \tau \le 1\big)$$

be the corresponding moment cone. Given a point $a = (\alpha_0, \dots, \alpha_d)$, we want to decide whether $a \in M$. In other words, (cf. Sections 16, 17 and 31), we want to decide whether there exists a Borel measure μ on the interval $[0, 1]$ with the prescribed moments:

$$\int_0^1 \tau^i \, d\mu = \alpha_i, \quad i = 1, \dots, d.$$

Let $K \subset \mathbb{R}^{d+1}$ be the cone of polynomials that are non-negative on $[0, 1]$ (see Section 18):

$$K = \Big\{(\gamma_0, \dots, \gamma_d) : \gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d \ge 0 : \quad \text{for all} \quad \tau \in [0,1]\Big\}.$$

(33.1) Lemma. *We have*

$$M = \Big\{x \in \mathbb{R}^{d+1} : \langle x, c\rangle \ge 0 \quad \textit{for all} \quad c \in K\Big\}.$$

Proof. We can write

$$\begin{aligned} K &= \Big\{c \in \mathbb{R}^{d+1} : \langle c, g(\tau)\rangle \ge 0 \quad \text{for all} \quad \tau \in [0,1]\Big\} \\ &= \Big\{c \in \mathbb{R}^{d+1} : \langle c, x\rangle \ge 0 \quad \text{for all} \quad x \in M\Big\}. \end{aligned}$$

Denoting $-K = \{-c : c \in K\}$ and using Lemma 32.4, we can write $-K = M^\circ$. Since both K and M are closed convex sets containing the origin (cf. Lemma 16.3), applying the Bipolar Theorem (Theorem 32.2), we get

$$M = (-K)^\circ = \Big\{x \in \mathbb{R}^{d+1} : \langle x, c\rangle \ge 0 \quad \text{for all} \quad c \in K\Big\}$$

and the proof follows. □

(33.2) Proposition. *Let $a = (\alpha_0, \dots, \alpha_d)$ be a point.*

(1) *Suppose that d is even, $d = 2m$. Then $a \in M$ if and only if*

$$\sum_{i,j=0}^{m} \alpha_{i+j}\xi_i\xi_j \ge 0 \quad \textit{for all} \quad x = (\xi_0, \dots, \xi_m) \in \mathbb{R}^{m+1}$$

and

$$\sum_{i,j=0}^{m-1} \big(\alpha_{i+j+1} - \alpha_{i+j+2}\big)\xi_i\xi_j \ge 0 \quad \textit{for all} \quad x = (\xi_0, \dots, \xi_{m-1}) \in \mathbb{R}^m.$$

(2) *Suppose that d is odd, $d = 2m+1$. Then $a \in M$ if and only if*

$$\sum_{i,j=0}^{m} \alpha_{i+j+1}\xi_i\xi_j \geq 0 \quad \textit{for all} \quad x = (\xi_0, \ldots, \xi_m) \in \mathbb{R}^{m+1}$$

and

$$\sum_{i,j=0}^{m} \big(\alpha_{i+j} - \alpha_{i+j+1}\big)\xi_i\xi_j \geq 0 \quad \textit{for all} \quad x = (\xi_0, \ldots, \xi_m) \in \mathbb{R}^{m+1}.$$

Proof. Suppose that $d = 2m$. Corollary 18.3 asserts that the polynomial $p(\tau) = \gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d$ is non-negative on the interval $[0,1]$ if and only if p can be written as a convex combination of polynomials q_i^2, where $q_i = \xi_0 + \xi_1\tau + \ldots + \xi_m\tau^m$ are polynomials of degree m, and polynomials $\tau(1-\tau)q_j^2$, where $q_j(\tau) = \xi_0 + \xi_1\tau + \ldots + \xi_{k-1}\tau^{m-1}$ is are polynomials of degree $m-1$.

Applying Lemma 33.1, we conclude that $a \in M$ if and only if

$\langle a, c\rangle \geq 0$ for all points $c = (\gamma_0, \ldots, \gamma_d)$ such that

$$\gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d = (\xi_0 + \xi_1\tau + \ldots + \xi_m\tau^m)^2$$

for some $x = (\xi_0, \ldots, \xi_m) \in \mathbb{R}^{m+1}$

and $\langle a, c\rangle \geq 0$ for all points $c = (\gamma_0, \ldots, \gamma_d)$ such that

$$\gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d = \tau(1-\tau)(\xi_0 + \xi_1\tau + \ldots + \xi_{m-1}\tau^{m-1})^2$$

for some $x = (\xi_1, \ldots, \xi_{m-1}) \in \mathbb{R}^{m-1}$.

In other words, $a \in M$ if and only if

$$\sum_{k=0}^{d} \alpha_k \Big(\sum_{(i,j):i+j=k} \xi_i\xi_j \Big) \geq 0 \quad \text{for all} \quad x = (\xi_0, \ldots, \xi_m) \quad \text{and}$$

$$\sum_{k=0}^{d} \alpha_k \Big(\sum_{(i,j):i+j=k-1} \xi_i\xi_j - \sum_{(i,j):i+j=k-2} \xi_i\xi_j \Big) \geq 0 \quad \text{for all} \quad x = (\xi_0, \ldots, \xi_{m-1}).$$

The first part now follows.

Suppose that $d = 2m+1$. Corollary 18.3 asserts that the polynomial $p(\tau) = \gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d$ is non-negative on the interval $[0,1]$ if and only if p can be written as a convex combination of polynomials τq_i^2 and $(1-\tau)q_j^2$, where $q_i, q_j = \xi_0 + \xi_1\tau + \ldots + \xi_m\tau^m$ are polynomials of degree m.

Applying Lemma 33.1, we conclude that $a \in M$ if and only if

$\langle c, a\rangle \geq 0$ for all points $c = (\gamma_0, \ldots, \gamma_d)$ such that

$$\gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d = \tau(\xi_0 + \xi_1\tau + \ldots + \gamma_m\tau^m)^2$$

for some $x = (\xi_0, \ldots, \xi_m) \in \mathbb{R}^{m+1}$

and $\langle c, a\rangle \geq 0$ for all points $c = (\gamma_0, \ldots, \gamma_d)$ such that

$$\gamma_0 + \gamma_1\tau + \ldots + \gamma_d\tau^d = (1-\tau)(\xi_0 + \xi_1\tau + \ldots + \gamma_m\tau^m)^2$$

for some $x = (\xi_0, \ldots, \xi_m) \in \mathbb{R}^{m+1}$.

In other words, $a \in M$ if and only if

$$\sum_{k=0}^{d} \alpha_k \Big(\sum_{(i,j):i+j=k-1} \xi_i\xi_j \Big) \geq 0 \quad \text{for all} \quad x = (\xi_0, \ldots, \xi_m) \quad \text{and}$$

$$\sum_{k=0}^{d} \alpha_k \Big(\sum_{(i,j):i+j=k} \xi_i\xi_j - \sum_{(i,j):i+j=k-1} \xi_i\xi_j \Big) \geq 0 \quad \text{for all} \quad x = (\xi_0, \ldots, \xi_m)$$

and the proof of Part 2 follows. □

(33.3) Corollary. *Let $a = (\alpha_0, \ldots, \alpha_d)$ be a point.*

(1) *Suppose that d is even, $d = 2m$.*

For $n = 0, \ldots, m$, let A_n be the $(n+1) \times (n+1)$ matrix whose (i,j)-th entry is α_{i+j-2}:

$$A_n = \begin{pmatrix} \alpha_0 & \alpha_1 & \ldots & \alpha_n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_n & \alpha_{n+1} & \ldots & \alpha_{2n} \end{pmatrix}.$$

For $n = 0, \ldots, m-1$ let A'_n be the $(n+1) \times (n+1)$ matrix whose (i,j)-th entry is $\alpha_{i+j-1} - \alpha_{i+j}$,

$$A'_n = \begin{pmatrix} \alpha_1 - \alpha_2 & \ldots & \alpha_{n+1} - \alpha_{n+2} \\ \alpha_2 - \alpha_3 & \ldots & \alpha_{n+2} - \alpha_{n+3} \\ \ldots & \ldots & \ldots \\ \alpha_{n+1} - \alpha_{n+2} & \ldots & \alpha_{2n+1} - \alpha_{2n+2} \end{pmatrix}.$$

Then $a \in M$ if and only if $\det A_n \geq 0$ for $n = 0, \ldots, m$ and $\det A'_n \geq 0$ for $n = 0, \ldots, m-1$.

(2) *Suppose that d is odd, $d = 2m+1$.*

For $n = 0, \dots, m$ let A_n be the $(n+1) \times (n+1)$ matrix whose (i,j)-th entry is α_{i+j-1},

$$A_n = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{n+1} \\ \alpha_2 & \alpha_3 & \dots & \alpha_{n+2} \\ \dots & \dots & \dots & \dots \\ \alpha_{n+1} & \alpha_{n+2} & \dots & \alpha_{2n+1} \end{pmatrix}.$$

For $n = 0, \dots, m$ let A'_n be the $(n+1) \times (n+1)$ matrix whose (i,j)-th entry is $\alpha_{i+j-2} - \alpha_{i+j-1}$,

$$A'_n = \begin{pmatrix} \alpha_0 - \alpha_1 & \dots & \alpha_n - \alpha_{n+1} \\ \alpha_1 - \alpha_2 & \dots & \alpha_{n+1} - \alpha_{n+2} \\ \dots & \dots & \dots \\ \alpha_n - \alpha_{n+1} & \dots & \alpha_{2n} - \alpha_{2n+1} \end{pmatrix}.$$

Then $a \in M$ if and only if $\det A_n \geq 0$ and $\det A'_n \geq 0$ for $n = 0, \dots, m$.

Proof. Follows from Proposition 33.2 by Sylvester's criterion, which gives necessary and sufficient conditions for a quadratic form to be positive semidefinite. □

PROBLEM.

1. Draw a picture of M and check the conditions of Corollary 33.3 when $d = 2$.

34. Polytopes and Polarity

Recall (Definition 2.2) that a polytope is the convex hull of a finite set of points in $\mathbb{R}^d$. We will discuss the facial structure of polytopes in some detail.

(34.1) Lemma. *Let $P = \operatorname{conv}\{v_1, \dots, v_m\} \subset \mathbb{R}^d$ be a polytope and let $F \subset P$ be its face. Then $F = \operatorname{conv}\{v_i : v_i \in F\}$. In particular, a face of a polytope is a polytope and the number of faces of a polytope is finite.*

Proof. Since F is a face of P, there exists a linear function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ and a number α such that $f(x) \leq \alpha$ for all $x \in P$ and $f(x) = \alpha$ if and only if $x \in F$ (see Definition 9.6). Let $I = \{i : f(v_i) = \alpha\}$, so $f(v_i) < \alpha$ for all $i \notin I$. Obviously, $\operatorname{conv}(v_i : i \in I) \subset F$. It remains to show that $F \subset \operatorname{conv}(v_i : i \in I)$. Let us choose an $x \in F$. Since $x \in P$, we have

$$x = \sum_{i=1}^{m} \lambda_i v_i, \quad \text{where} \quad \lambda_i \geq 0 \quad \text{for} \quad i = 1, \dots, m \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i = 1.$$

If $\lambda_i > 0$ for some $i \notin I$, then $f(\lambda_i v_i) = \lambda_i f(v_i) < \lambda_i \alpha$ and $f(x) = \sum_{i=1}^{m} \lambda_i f(v_i) < \alpha \sum_{i=1}^{m} \lambda_i = \alpha$, which is a contradiction. Hence $\lambda_i = 0$ for $i \notin I$ and $x \in \operatorname{conv}(v_i : i \in I)$. □

(34.2) Corollary. *Let P be a polytope, let F be a face of P and let G be a face of F. Then G is a face of P.*

Proof. Suppose that $P = \operatorname{conv}\big(v_1, \ldots, v_m\big)$, $F = \operatorname{conv}\big(v_i : i \in I\big)$ and $G = \operatorname{conv}\big(v_j : j \in J \subset I\big)$. There exists a linear function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ and a number α such that $f(v_i) = \alpha$ for all $i \in I$ and $f(v_i) < \alpha$ for all $i \notin I$, and there exists a linear function $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ and a number β such that $g(v_i) = \beta$ for $i \in J$ and $g(v_i) < \beta$ for $i \in I \setminus J$. Then, for a sufficiently small $\epsilon > 0$, the function $h = f + \epsilon g$ has the following property: $h(v_i) = \alpha + \epsilon\beta$ for $i \in J$ and $h(v_i) < \alpha + \epsilon\beta$ for $i \notin J$. Hence G is a face of P. $\square$

PROBLEMS.

1°. Prove that an extreme point of a polytope is a face (exposed point) of the polytope.

2°. Prove that a face of a polyhedron is a polyhedron.

3. Let $P = \big\{x \in \mathbb{R}^d : \langle c_i, x\rangle \le \beta_i, i = 1, \ldots, m\big\}$ be a polyhedron and let $u \in P$ be a point. Let $I_u = \big\{i : \langle c_i, u\rangle = \beta_i\big\}$. Let $F = \big\{x \in P : \langle c_i, x\rangle = \beta_i : i \in I_u\big\}$. Prove that F is the smallest face of P containing x (we agree here that P is a face of itself).

4. Prove that a polyhedron has finitely many faces.

5. Let $P \subset \mathbb{R}^d$ be a polyhedron, let $F \subset P$ be a face of P and let $G \subset F$ be a face of F. Prove that G is a face of P.

6°. Prove that the intersection of two polytopes is a polytope.

Now is the main result of this section, which establishes an inclusion reversing correspondence between the faces of a polytope and the faces of its polar.

(34.3) Theorem. *Let $P \subset \mathbb{R}^d$ be a polytope, containing the origin in its interior and let $Q = P^\circ$. Then Q is a polytope. For a face F of P, let*

$$\widehat{F} = \Big\{x \in Q : \langle x, y\rangle = 1 \quad \textit{for each} \quad y \in F\Big\}$$

(we agree that $\widehat{\emptyset} = Q$ and that $\widehat{P} = \emptyset$). Then

(1) *The set $\widehat{F}$ is a face of Q and*

$$\dim F + \dim \widehat{F} = d - 1;$$

(2) *If $F \subset G$ are faces of P then $\widehat{G} \subset \widehat{F}$;*

(3) *Let G be a face of Q and let*

$$F = \Big\{y \in P : \langle x, y\rangle = 1 \quad \textit{for each} \quad x \in G\Big\}$$

(we agree that $F = P$ for $G = \emptyset$ and that $F = \emptyset$ for $G = Q$). Then

$$\widehat{F} = G.$$

Proof. The proof that Q is a polytope is incorporated in the proof of Corollary 32.3. Suppose that $P = \operatorname{conv}\big(v_1, \dots, v_m\big)$. By Lemma 34.1, $F = \operatorname{conv}\big(v_i : i \in I\big)$ for some set of indices I. Let $v = \dfrac{1}{|I|} \sum_{i \in I} v_i$, where $|I|$ is the cardinality of I. Hence $v \in F$. We claim that

$$\widehat{F} = \big\{ x \in Q : \langle x, v \rangle = 1 \big\}. \tag{34.3.1}$$

Indeed, $v \in F$, so $\langle x, v \rangle = 1$ for any $x \in \widehat{F}$. On the other hand, for any $x \in Q$ and any v_i, we have $\langle x, v_i \rangle \le 1$. Therefore, if $x \in Q$ and $\langle x, v \rangle = 1$, then $\langle x, v_i \rangle = 1$ for $i \in I$ and hence $x \in \widehat{F}$. By (34.3.1), we conclude that $\widehat{F}$ is a face of Q.

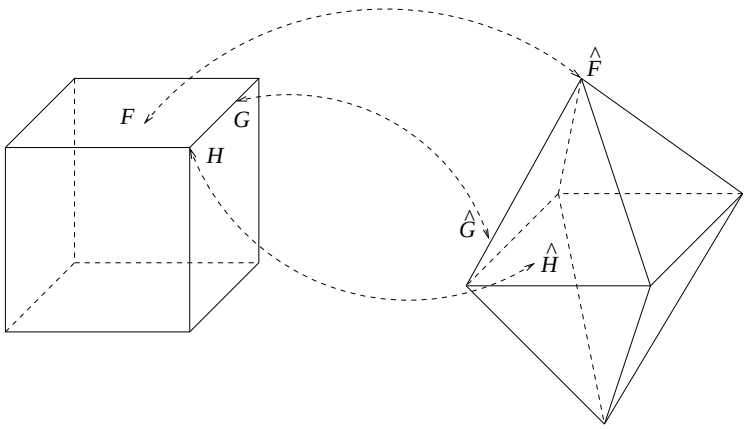
Let us prove that $\dim F + \dim \widehat{F} = d - 1$. Since F is a face, there is a vector $c \in \mathbb{R}^d$ and a number α, such that

$$\langle c, v_i \rangle = \alpha \quad \text{for} \quad i \in I \qquad \text{and} \qquad \langle c, v_i \rangle < \alpha \quad \text{for} \quad i \notin I.$$

Since $0 \in \operatorname{int} P$, $\alpha > 0$. Scaling, if necessary, we can assume that $\alpha = 1$. Then $c \in \widehat{F}$. Suppose that $\dim F = k$. For any $x \in \widehat{F}$, we must have $\langle x, v_i \rangle = 1$ for $i \in I$, or, in other words, $\langle x - c, v_i \rangle = 0$ for $i \in I$. Since $\dim \operatorname{span}(F) = k + 1$, we get that $\dim \widehat{F} \le d - k - 1$. On the other hand, if $y \in \mathbb{R}^d$ is a vector such that $\langle y, v_i \rangle = 0$ for $i \in I$, then $x = c + \epsilon y \in \widehat{F}$ for a sufficiently small $\epsilon > 0$. Hence $\dim \widehat{F} = d - k - 1$. Part 1 is proved.

Let us prove Part 2 of the Theorem. For any $x \in \widehat{G}$, we have $\langle x, y \rangle = 1$ for any $y \in G$. Since $F \subset G$, we have $\langle x, y \rangle = 1$ for any $y \in F$. Therefore, $x \in \widehat{F}$ and the proof of Part 2 follows.

By Theorem 32.2, $Q^\circ = P$. Then, by Part 1, F is a face of P and it is clear that $G \subset \widehat{F}$. Let $c \in \mathbb{R}^d$ be a vector such that $\langle c, x \rangle = 1$ for all $x \in G$ and $\langle c, x \rangle < 1$ for $x \notin Q \setminus G$ (see the proof of Part 2 above). Then $c \in F$. Therefore, for all $x \in \widehat{F}$ we have $\langle x, c \rangle = 1$, so $x \in G$. Therefore, $\widehat{F} \subset G$. Hence $G = \widehat{F}$ and the proof of Part 3 follows. $\square$



The polarity correspondence between faces

Figure 46

(34.4) Definitions. A 0-dimensional face of a polytope is called a *vertex*. A 1-dimensional face of a polytope is called an *edge*. A $(d-1)$-dimensional face of a d-dimensional polytope is called a *facet*. A $(d-2)$-dimensional face of a d-dimensional polytope is called a *ridge*.

PROBLEMS.

1. Prove that a d-dimensional polytope has a facet. Deduce that a d-dimensional polytope has a k-dimensional face for each $0 \leq k \leq d-1$.

2. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope. Prove that every ridge of P belongs to precisely two facets.

Hint: What is the dual statement?

3. Prove the following *diamond property*. Let P be a polytope and $G \subset F$ be two faces of P such that $\dim F - \dim G = 2$. Then there are precisely two faces H_1, H_2, such that $G \subset H_1, H_2 \subset F$ (all inclusions are proper).

4. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope containing the origin in its interior and let $F \subset P$ be a k-dimensional face of P and let $\widehat{F} \subset P^\circ$ be the face from Theorem 34.3. Hence F is a $(d-k-1)$-dimensional polytope. We consider F as a full-dimensional polytope in its affine hull and choose the origin in it interior. Let $R = F^\circ$. Hence R is a $(d-k-1)$-dimensional polytope. Establish an inclusion preserving bijection ϕ between the faces of R and the faces H of P containing F. The (facial structure of) the polytope R is called the *face figure* of F in P and

denoted P/F.

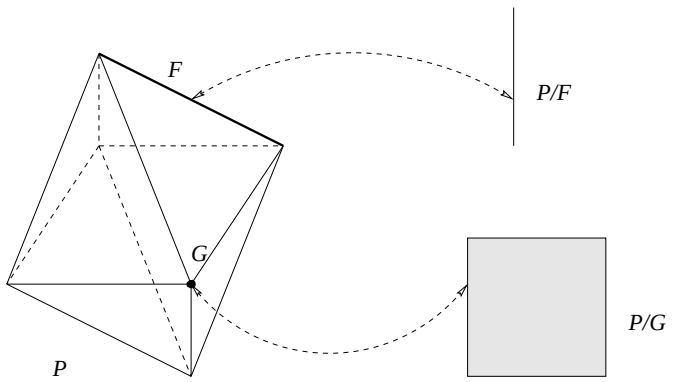


Figure 47

5. Let $K \subset \mathbb{R}^d$ be a convex compact set containing the origin in its interior. Let $Q_0 = K^\circ$ and let Q_p be the polar of K with the origin moved to p: $Q_p = \{x \in \mathbb{R}^d : \langle x, y - p\rangle \leq 1 \text{ for any } y \in K\}$. Prove, that Q_p is the image of Q_0 under the *projective transformation*:

$$x \longmapsto \frac{x}{1 - \langle x, p\rangle}.$$

We will need the following useful result, which can be considered as a sharpening of Corollary 32.3.

(34.5) Lemma. *Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope. Then P can be represented in the form*

$$P = \Big\{x \in \mathbb{R}^d : l_i(x) \leq \alpha_i, \quad i = 1, \ldots, m\Big\},$$

where $l_i : \mathbb{R}^d \longrightarrow \mathbb{R}$ are linear functions, $\alpha_i \in \mathbb{R}$ are numbers and the sets

$$F_i = \Big\{x \in P : l_i(x) = \alpha_i\Big\}$$

are the facets of P for $i = 1, \ldots, m$.

Proof. The proof follows the proof of Corollary 32.3 with some modifications. Let us choose the origin in the interior of P and let $Q = P^\circ$ be the polar of P. We can write $Q = \operatorname{conv}(v_i : i = 1, \ldots, m)$, where v_i are the extreme points of Q. By Problem 1 of Section 34.2, v_i are the vertices of Q. Then $P = \{x : \langle v_i, x\rangle \leq 1 : i = 1, \ldots, m\}$. Let us choose $f_i(x) = \langle v_i, x\rangle$ and $\alpha_i = 1$. By Theorem 34.3, $F_i = \widehat{v_i}$ are the facets of P. □

35. The Facial Structure of the Permutation Polytope

Recall (see Definition 13.1) that the permutation polytope $P(a)$, where $a = (\alpha_1, \ldots, \alpha_n)$, $a \in \mathbb{R}^n$, is the convex hull of the points $\sigma(a)$ obtained from a by permutations of the coordinates. In this section, we describe the facial structure of $P(a)$ assuming that the coordinates $\alpha_1, \ldots, \alpha_n$ are distinct. We need a simple, yet useful result.

(35.1) Lemma. *Let $x = (\xi_1, \ldots, \xi_n)$ and $y = (\eta_1, \ldots, \eta_n)$ be n-vectors. Suppose that $\xi_i > \xi_j$ and $\eta_i < \eta_j$. Let $\overline{y}$ be the vector obtained from y by swapping η_i and η_j. Then*

$$\langle x, \overline{y} \rangle > \langle x, y \rangle.$$

Proof. We have

$$\langle x, \overline{y} \rangle - \langle x, y \rangle = \xi_i \eta_j + \xi_j \eta_i - \xi_i \eta_i - \xi_j \eta_j = (\xi_i - \xi_j)(\eta_j - \eta_i) > 0.$$

□

PROBLEM.

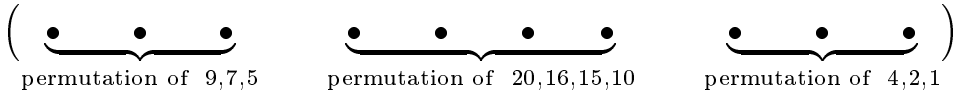
1°. Suppose that there are 6 boxes in front of you, the first is stuffed with \$ 100 bills, the second with \$ 50 bills, the third with \$ 20 bills, the fourth with \$ 10 bills, the fifth with \$ 5 bills and the sixth with \$ 1 bills. You are allowed to take from the i-th box a number β_i of bills with the only condition that the (unordered) set of numbers $\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\}$ is the (unordered) set $\{3, 5, 10, 2, 15, 8\}$. How many bills should you take from each box to maximize the total amount taken?

(35.2) Theorem. *Let $a = (\alpha_1, \ldots, \alpha_n)$ be a point such that $\alpha_1 > \alpha_2 > \ldots > \alpha_n$ and let $P = P(a)$ be the corresponding permutation polytope. For a number $1 \leq k \leq n$, let $\mathcal{S}$ be a partition of the set $\{1, \ldots, n\}$ into k pairwise disjoint non-empty subsets $S_1, \ldots, S_k$. Let $s_i = |S_i|$ be the cardinality of the i-th subset for $i = 1, \ldots, k$, let $t_i = \sum_{j=1}^{i} s_j$ for $i = 1, \ldots, k$ and let us define sets $A_1 = \{\alpha_j : 1 \leq j \leq s_1\}$ and $A_i = \{\alpha_j : t_{i-1} \leq j \leq t_i\}$ for $i = 2, \ldots, k$.*

Let $F_{\mathcal{S}}$ be the convex hull of the points $b = \sigma(a)$, $b = (\beta_1, \ldots, \beta_n)$ such that $\{\beta_j : j \in S_i\} = A_i$ for all $i = 1, \ldots, k$.

Then $F_{\mathcal{S}}$ is a face of P, $\dim F_{\mathcal{S}} = n - k$ and for every face F of P we have $F = F_{\mathcal{S}}$ for some partition $\mathcal{S}$.

Example. Let $n = 10$, $a = (20, 16, 15, 10, 9, 7, 5, 4, 2, 1)$, $S_1 = \{4, 5, 6, 7\}$, $S_2 = \{1, 2, 3\}$ and $S_3 = \{8, 9, 10\}$. Hence $A_1 = \{20, 16, 15, 10\}$, $A_2 = \{9, 7, 5\}$ and $A_3 = \{4, 2, 1\}$. The face $F_{\mathcal{S}}$ is the convex hull of all points $b = \sigma(a)$ of the type



and $\dim F_{\mathcal{S}} = 7$.

Proof of Theorem 35.2. Let us describe all the faces F of P containing a.

Let $c = (\gamma_1, \ldots, \gamma_n)$ be a vector and λ be a number such that $\langle c, x\rangle \le \lambda$ for all $x \in P$ and $\langle c, x\rangle = \lambda$ if and only if $x \in F$. Since $a \in F$, we have $\langle c, a\rangle = \lambda$. Lemma 35.1 implies that we must have $\gamma_1 \ge \gamma_2 \ge \ldots \ge \gamma_n$, since if for some $i < j$ we had $\gamma_i < \gamma_j$, we would have obtained $\langle c, \sigma(a)\rangle > \langle c, a\rangle$ for the transposition σ that swaps α_i and α_j.

Let us split the sequence $\gamma_1 \ge \gamma_2 \ge \ldots \ge \gamma_n$ into the subintervals $S_1, \ldots, S_k$ for which γ's do not change:

$$\underbrace{\gamma_1 = \ldots = \gamma_{t_1}}_{S_1=\{1,\ldots,t_1\}} > \underbrace{\gamma_{t_1+1} = \ldots = \gamma_{t_2}}_{S_2=\{t_1+1,\ldots,t_2\}} > \gamma_{t_2+1} = \ldots > \underbrace{\gamma_{t_{k-1}+1} = \ldots = \gamma_n}_{S_k=\{t_{k-1}+1,\ldots,n\}}.$$

Hence $S_1 = \{j : \gamma_j = \gamma_1\}$, $s_1 = t_1 = |S_1|$ and $S_i = \{j : \gamma_j = \gamma_{t_{i-1}+1}\}$, $s_i = |S_i|$ and $t_i = t_{i-1} + s_i$ for $i = 2, \ldots, k$.

We observe that for $b = \sigma(a)$, $b = (\beta_1, \ldots, \beta_n)$, we have $\langle b, c\rangle = \langle a, c\rangle$ if and only if $(\beta_1, \ldots, \beta_{t_1})$ is a permutation of $(\alpha_1, \ldots, \alpha_{t_1})$, $(\beta_{t_1+1}, \ldots, \beta_{t_2})$ is a permutation of $(\alpha_{t_1+1}, \ldots, \alpha_{t_2})$, and so forth, so that $(\beta_{t_{k-1}+1}, \ldots, \beta_n)$ is a permutation of $(\alpha_{t_{k-1}+1}, \ldots, \alpha_n)$. Applying Lemma 34.1, we conclude that $F = F_{\mathcal{S}}$ for the partition $\mathcal{S} = \{S_1, \ldots, S_k\}$. In this case, we have $A_1 = \{\alpha_1, \ldots, \alpha_{s_1}\}$ and $A_i = \{\alpha_{t_{i-1}+1}, \ldots, \alpha_{t_i}\}$ for $i = 2, \ldots, k$.

Vice versa, every vector $c = (\gamma_1, \ldots, \gamma_n)$ with $\gamma_1 \ge \gamma_2 \ge \ldots \ge \gamma_n$ gives rise to a face $F_{\mathcal{S}}$ containing a, where $\mathcal{S} = S_1 \cup S_2 \cup \ldots \cup S_k$ is the partition of $\{1, \ldots, n\}$ into the subintervals on which γ's do not change.

Let $a_1 = (\alpha_1, \ldots, \alpha_{s_1}) \in \mathbb{R}^{s_1}$ and let $a_i = (\alpha_{t_{i-1}+1}, \ldots, \alpha_{t_i}) \in \mathbb{R}^{s_i}$ for $i = 2, \ldots, k$. Geometrically, the face $F_{\mathcal{S}}$ is the direct product

$$F_{\mathcal{S}} = P(a_1) \times \ldots \times P(a_k)$$

of the permutation polytopes $P(a_i) \subset \mathbb{R}^{s_i}$. Since a has distinct coordinates, the coordinates of each a_i are distinct as well, so by Problem 4, Section 13.1, we have $\dim P(a_i) = s_i - 1$ for $i = 1, \ldots, k$. Therefore,

$$\dim F_{\mathcal{S}} = \sum_{i=1}^{k} \dim P(a_i) = \sum_{i=1}^{k} (s_i - 1) = \left(\sum_{i=1}^{k} s_i\right) - k = n - k.$$

Hence we have described the faces F of P containing a. Let $\sigma \in S_n$ be a permutation and let $\sigma(x) = y$ for $x = (\xi_1, \ldots, \xi_n)$ and $y = (\eta_1, \ldots, \eta_n)$ provided $\eta_i = \xi_{\sigma^{-1}(i)}$. Then the action $x \longmapsto \sigma(x)$ is an orthogonal transformation of $\mathbb{R}^n$. Hence we conclude that F is a face of P if and only if for some permutation σ, the set $\sigma(F)$ is a face of P containing a. If $\sigma(F) = F_{\mathcal{S}}$ for $\mathcal{S} = S_1 \cup \ldots \cup S_k$, then $F = F_{\mathcal{S}'}$, where $\mathcal{S}' = \sigma^{-1}(S_1) \cup \ldots \cup \sigma^{-1}(S_k)$. This completes the proof. □

PROBLEMS.

1°. Let $a = (\alpha_1, \ldots, \alpha_n)$ be a vector such that $\alpha_1 > \alpha_2 > \ldots > \alpha_n$ and let $P(a)$ be the corresponding permutation polytope. Let $b = (\beta_1, \ldots, \beta_n)$ be a vertex of P. Let b' be another vertex of P. Prove that the interval $[b, b']$ is an edge of P if and only if b' is obtained from b by swapping two values β_i and β_j such that $\beta_i = \alpha_k$ and $\beta_j = \alpha_{k+1}$ for some $k = 1, \ldots, n-1$.

2. Let $a \in \mathbb{R}^n$ be a vector with distinct components and let $P(a)$ be the corresponding permutation polytope. Let f_k be the number of k-dimensional faces of P. Prove that

$$f_k = \sum_{\substack{m_1+\ldots+m_{n-k}=n \\ m_1,\ldots,m_{n-k} \text{ are positive integers}}} \frac{n!}{m_1! \cdots m_{n-k}!}.$$

3*. Let us choose $a = (n-1, n-2, \ldots, 1, 0)$, $a \in \mathbb{R}^n$. The polytope $P_{n-1} = P(a)$ is called the *permutohedron*. Prove that the permutohedron is a *zonotope*, that is the Minkowski sum of finitely many (namely, $\binom{n}{2}$) straight line intervals.

Hint: Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. For $i < j$ let

$$I_{ij} = \operatorname{conv}\Big(\frac{1}{2}e_i - \frac{1}{2}e_j, \quad \frac{1}{2}e_j - \frac{1}{2}e_i\Big).$$

Then

$$P_{n-1} = \frac{n-1}{2}(1, \ldots, 1) + \sum_{i<j} I_{ij},$$

where "$\sum$" stands for the Minkowski sum. See Example 7.15 of G.M. Ziegler, *Lectures on Polytopes*, Springer, 1995.

As an corollary of Theorem 35.2, we obtain a theorem of R. Rado (1952), which gives us a useful criterion to check whether a given point belongs to the given permutation polytope.

(35.3) Rado's Theorem. *Let us $a = (\alpha_1, \ldots, \alpha_n)$ be a point, where $\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$ and let $P(a) \subset \mathbb{R}^n$ be the corresponding permutation polytope. Let $b \in \mathbb{R}^n$ be a point and let $\beta_1 \geq \beta_2 \geq \ldots \geq \beta_n$ be the ordering of the coordinates of b. Then $b \in P(a)$ if and only if*

$$\sum_{i=1}^{k} \beta_i \leq \sum_{i=1}^{k} \alpha_i \quad \textit{for} \quad k = 1, \ldots, n-1$$

and

$$\sum_{i=1}^{n} \beta_i = \sum_{i=1}^{n} \alpha_i.$$

Proof. We prove the result assuming that $\alpha_1 > \alpha_2 > \ldots > \alpha_n$. The proof in the general case follows then by a continuity argument.

We know that $\dim P(a) = n-1$ and that $P(a)$ lies in the affine hyperplane

$$H = \Big\{(\xi_1, \dots, \xi_n) : \xi_1 + \dots + \xi_n = \alpha_1 + \dots + \alpha_n\Big\}$$

(see Problem 3, Section 13.1). Theorem 35.2 implies that the facets $F_{\mathcal{S}}$ of P are indexed by the partitions $\mathcal{S} = S_1 \cup S_2$ of the set $\{1, \dots, n\}$ into non-empty disjoint subsets. Suppose that $|S_1| = s_1$ and $|S_2| = s_2$, so $s_1, s_2 > 0$ and $s_1 + s_2 = n$. The facet $F_{\mathcal{S}}$ is a convex hull of the points $x = \sigma(a)$, $x = (\xi_1, \dots, \xi_n)$ such that

$$\big\{\xi_i : i \in S_1\big\} = \big\{\alpha_1, \dots, \alpha_{s_1}\big\} \quad \text{and} \quad \big\{\xi_i : i \in S_2\big\} = \big\{\alpha_{s_1+1}, \dots, \alpha_{s_n}\big\}.$$

It follows then that in H the facet $F_{\mathcal{S}}$ is defined by the inequality

$$\sum_{i \in S_1} \xi_i \le \sum_{i=1}^{s_1} \alpha_i \tag{35.3.1}$$

since (35.3.1) is trivially satisfied on all vertices of P and is sharp on the vertices from $F_{\mathcal{S}}$. Lemma 34.5 implies that $b \in P(a)$ if and only if $b \in H$ and b satisfies the inequality (35.3.1) for every facet $F_{\mathcal{S}}$ of P. Since for every subset $S_1 \subset \{1, \dots, n\}$ we have

$$\sum_{i \in S_1} \beta_i \le \sum_{i=1}^{s_1} \beta_i,$$

the proof follows. □

36. The Euler-Poincaré Formula

In this section, we prove a classical relation for the number of faces of an arbitrary polytope. Throughout this section, we use the Euler characteristic μ (see Section 6). For a set $A \subset \mathbb{R}^d$, we write simply $\mu(A)$ instead of $\mu([A])$. In what follows, we agree that a polytope P is a face of itself. We agree that the dimension of the empty face is -1.

(36.1) Lemma. *Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope and let ∂P be the boundary of P. Then*

$$\mu(\partial P) = 1 + (-1)^{d-1}.$$

Proof. Since ∂P is the union of faces of P, the indicator function $[\partial P]$ belongs to the algebra $\mathcal{K}(\mathbb{R}^d)$ of convex compact sets, so $\mu(\partial P)$ is defined. We prove the identity by induction on d. If $d = 1$ then P is an interval, so ∂P consists of two points and $\mu(\partial P) = 2$.

Let us introduce a family of hyperplanes $H_\tau = \big\{(\xi_1, \ldots, \xi_d) : \xi_d = \tau\big\}$, $\tau \in \mathbb{R}$.

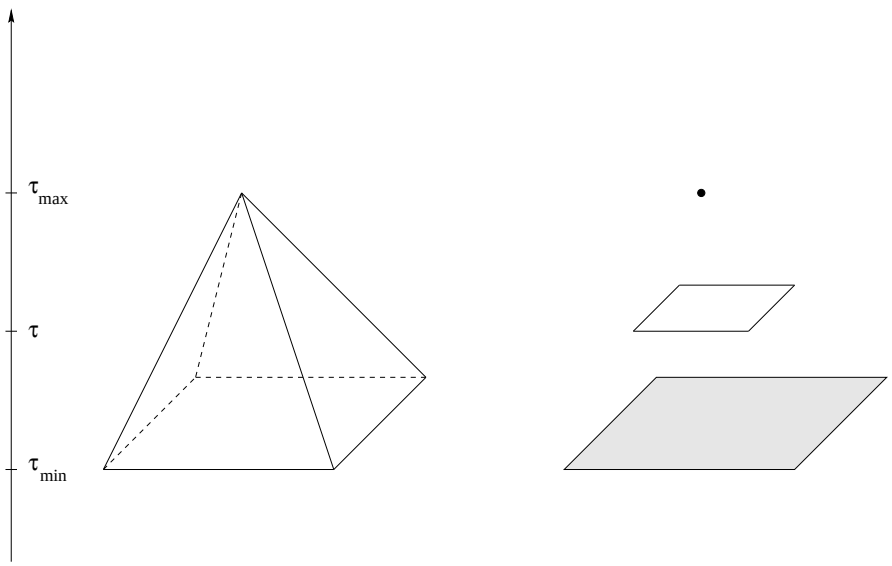


Figure 48

By Corollary 6.5,

$$\mu(\partial P) = \sum_{\tau \in \mathbb{R}} \Big(\mu(\partial P \cap H_\tau) - \lim_{\epsilon \longrightarrow +0} \mu(\partial P \cap H_{\tau - \epsilon}) \Big).$$

Let $\tau_{\min} = \min_{x \in P} \xi_d$ and let $\tau_{\max} = \max_{x \in P} \xi_d$. For any $\tau_{\min} < \tau < \tau_{\max}$, the intersection $P \cap H_\tau$ is a $(d-1)$-dimensional polytope and $\partial(P \cap H_\tau) = \partial P \cap H_\tau$. Therefore, $\mu(\partial P \cap H_\tau) = 1 + (-1)^{d-2}$ for all $\tau_{\min} < \tau < \tau_{\max}$. Furthermore, for $\tau = \tau_{\min}$ and $\tau = \tau_{\max}$, the intersection $P \cap H_\tau$ is a face of P lying in ∂P. Therefore, $\mu(\partial P \cap H_\tau) = 1$ for $\tau = \tau_{\max}$ or $\tau = \tau_{\min}$. Finally, for $\tau > \tau_{\max}$ and $\tau < \tau_{\min}$, the intersection $P \cap H_\tau$ is empty, so $\mu(P \cap H_\tau) = 0$. Summarizing,

$$\mu(\partial P) = 1 - (1 + (-1)^{d-2}) + 1 = 1 + (-1)^{d-1},$$

and the proof follows. □

(36.2) Definitions. A finite set $\mathcal{P} = \big\{P_i \subset \mathbb{R}^d, i \in I\big\}$ of distinct polytopes in $\mathbb{R}^d$ is called a *polytopal complex*, provided the following two conditions are satisfied:

(36.2.1) For any two polytopes $P_i, P_j \in \mathcal{P}$, the intersection $P_i \cap P_j$ is a face of both P_i and P_j

and

(36.2.2) If $P_i \in \mathcal{P}$ and F is a face of P_i, then $F \in \mathcal{P}$.

The union

$$|\mathcal{P}| = \bigcup_{i \in I} P_i$$

is called the *realization* of $\mathcal{P}$. The maximum dimension of a polytope from $\mathcal{P}$

$$k = \max_{P \in \mathcal{P}} \dim P$$

is called the *dimension* of $\mathcal{P}$ and denoted $\dim \mathcal{P}$.

PROBLEM.

1°. Let $\mathcal{P}$ be a polytopal complex and let $P \in \mathcal{P}$ be a polytope such that $\dim P = \dim \mathcal{P}$. Prove that $\mathcal{P} \setminus \{P\}$ is a polytopal complex.

(36.3) Theorem. *Let $\mathcal{P}$ be a polytopal complex, let $|\mathcal{P}| \subset \mathbb{R}^d$ be its realization and let $f_i(\mathcal{P})$ be the number of i-dimensional polytopes in $\mathcal{P}$. Then*

$$\mu(|\mathcal{P}|) = \sum_{i=0}^{d} (-1)^i f_i(\mathcal{P}).$$

Proof. We use induction on $k = \dim \mathcal{P}$. If $k = 0$, then $\mathcal{P}$ is a finite set of points from $\mathbb{R}^d$ and the formula is obvious.

Suppose that $k > 0$. Now we use induction on the number n of k-dimensional polytopes in $\mathcal{P}$.

Suppose that $n = 0$. Then $\dim \mathcal{P} < k$ and we apply the induction hypothesis.

Suppose that $n > 0$. Let us choose a polytope $P \in \mathcal{P}$ with $\dim P = k$. Let $\mathcal{P}_0 = \mathcal{P} \setminus P$. Then $\mathcal{P}_0$ is a polytopal complex (cf. Problem 1 of Section 36.2) and $\mathcal{P}_0$ contains $(n-1)$ polytope of the highest dimension k.

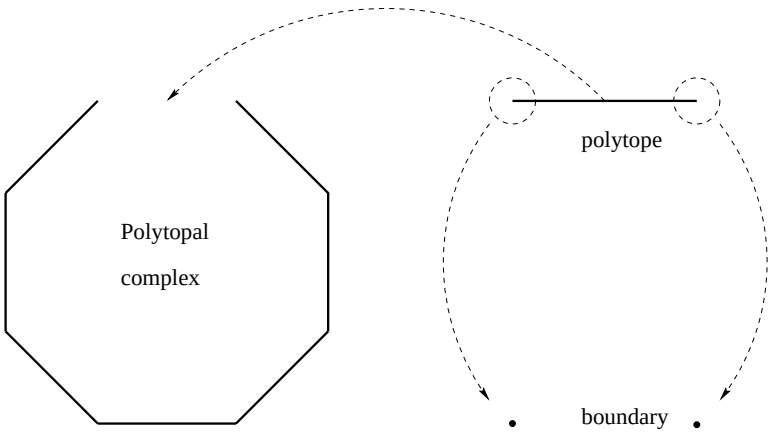


Figure 49

We can write

$$\mu\big(|\mathcal{P}|\big) = \mu\big(|\mathcal{P}_0|\big) + \mu(P) - \mu\big(|\mathcal{P}_0| \cap P\big).$$

We have

$$\mu\big(|\mathcal{P}_0|\big) = \sum_{i=0}^{d} (-1)^i f_i(\mathcal{P}_0)$$

by the induction hypothesis. Next, $\mu(P) = 1$, since P is a non-empty convex closed set. Finally, by (36.2.2), the intersection $|\mathcal{P}_0| \cap P$ is the boundary ∂P, considered in the affine hull of P. Applying Lemma 36.1, we conclude that $\mu\big(|\mathcal{P}_0| \cap P\big) = \mu(\partial P) = 1 + (-1)^{k-1}$. Summarizing,

$$\mu\big(|\mathcal{P}|\big) = \sum_{i=0}^{d} (-1)^i f_i(\mathcal{P}_0) + 1 - (1 + (-1)^{k-1}) = \sum_{i=0}^{d} (-1)^i f_i(\mathcal{P}_0) + (-1)^k.$$

We observe that $f_i(\mathcal{P}) = f_i(\mathcal{P}_0)$ for $i \neq k$ and that $f_k(\mathcal{P}) = f_k(\mathcal{P}_0) + 1$. The proof now follows. □

(36.4) Corollary (The Euler-Poincaré Formula). *Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope and let $f_i(P)$ be the number of i-dimensional faces of P. Then*

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

Proof. From Problem 5, Section 9.6 and Corollary 34.2, we conclude that the set of all proper faces of a polytope is a polytopal complex. The result follows by Theorem 36.3, applied to the polytopal complex consisting of the proper faces of P, and Lemma 36.1. □

The first complete proof of Corollary 36.4 was, apparently, obtained by H. Poincaré in 1893 (using methods of algebraic topology). A somewhat incomplete proof was given by L. Schläfli in 1852. The formula for $d \leq 3$ was known to L. Euler.

(36.5) Definition. Let P be a d-dimensional polytope.
The d-tuple $\big(f_0(P), \ldots, f_{d-1}(P)\big)$, where $f_i(P)$ is the number of i-dimensional faces of P, is called the *f-vector* of P. We often agree that $f_{-1} = f_d(P) = 1$.

PROBLEMS.

1. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope and let $H \subset \mathbb{R}^d$ be an affine hyperplane passing through an interior point of P and not containing any of the vertices of P. Let H_+ be an open halfspace bounded by H and let f_i^+ be the number of i-dimensional faces contained in H_+. Prove that

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

2. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope and let $l : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function, which takes different values on different vertices of P. For a vertex v of P, let us denote by f_i^v the number of i-dimensional faces F of P such that $l(v) = \max\{f(x) : x \in F\}$. Prove that

$$\sum_{i=0}^{d-1}(-1)^i f_i^v = \begin{cases} 1 & \text{if } l(v) \text{ is the minimum of } l \text{ on } P \\ (-1)^{d-1} & \text{if } l(v) \text{ is the maximum of } l \text{ on } P \\ 0 & \text{otherwise.} \end{cases}$$

3. Let P be a d-dimensional polytope and let $F \subset P$ be its k-dimensional face. Let $f_j(F, P)$ denote the number of j-dimensional faces of P that contain F. Prove that

$$\sum_{j=k}^{d-1}(-1)^j f_j(F, P) = (-1)^{d-1}.$$

4. Let $P \subset \mathbb{R}^d$ be an unbounded d-dimensional polyhedron, which does not contain a straight line. Let $f_i^0(P)$ be the number of bounded i-dimensional faces of P, let f_i^∞ be the number of unbounded i-dimensional faces of P and let $f_i(P) = f_i^0(P) + f_i^\infty(P)$ be the total number of i-dimensional faces of P. Prove that

$$\sum_{i=0}^{d-1}(-1)^i f_i^0(P) = 1, \quad \sum_{i=1}^{d}(-1)^{i+1} f_i^\infty(P) = 1 \quad \text{and} \quad \sum_{i=0}^{d}(-1)^i f_i(P) = 0.$$

5. Let $P \subset \mathbb{R}^d$ be a 3-dimensional polytope. Let us define the *curvature* $\kappa(v)$ at a vertex v of P as follows. Let $F_1, \ldots, F_m$ be the facets containing v. Then F_i is a planar polygon and let α_i be the angle at the vertex v of F_i. Let

$$\kappa(v) = 2\pi - \sum_{i=1}^{m} \alpha_i.$$

Prove the *Gauss-Bonnet formula*:

$$\sum_v \kappa(v) = 4\pi,$$

where the sum is taken over all vertices v of P.

The general Gauss-Bonnet formula asserts that the value of the analogous sum taken over the vertices of an oriented polyhedral 2-dimensional surface S is $2\pi\mu(S)$, where μ is the Euler characteristic. For smooth surfaces the sum is replaced by an integral.

6. Let us fix $d \geq 1$. In $\mathbb{R}^d$, consider the set of all f-vectors of d-dimensional polytopes. Prove that this set spans the hyperplane $\sum_{i=0}^{d-1}(-1)^i f_i = 1 + (-1)^{d-1}$. In other words, prove that there are no linear relations for the numbers f_i of i-dimensional faces other than the Euler-Poincaré formula, which would hold for all d-dimensional polytopes.

37. Polytopes with Many Faces: Cyclic Polytopes

Our goal is to construct polytopes with many faces. To this end, we modify the definition of the moment curve, see Section 16.

(37.1) Definition. Let us fix an interval, say, $[0,1] \subset \mathbb{R}^1$. Let

$$q(\tau) = (\tau, \tau^2, \ldots, \tau^d): \quad 0 \leq \tau \leq 1$$

be a curve in $\mathbb{R}^d$. Let us pick n distinct points $0 < \tau_1 < \tau_2 < \ldots < \tau_n < 1$ and let $v_i = q(\tau_i)$ for $i = 1, \ldots, n$. The polytope

$$C(d,n) = \operatorname{conv}(v_1, \ldots, v_n)$$

is called the *cyclic polytope*.

Although the construction depends on how the points $\tau_1 < \tau_2 < \ldots < \tau_n$ are chosen, we will see soon that the facial structure of $C(d,n)$ is independent of the choice.

PROBLEMS.

1°. Prove that each affine hyperplane $H \subset \mathbb{R}^d$ intersects the moment curve $q(\tau)$ in at most d points.

2°. Prove that $\dim C(d,n) = d$ for $n \geq d+1$.

(37.2) Proposition. *Let $C(d,n) = \operatorname{conv}(v_1, \ldots, v_n)$ be a cyclic polytope. Let $I \subset \{1, \ldots, n\}$ be a set such that $|I| \leq d/2$. Then $F = \operatorname{conv}\big(v_i : i \in I\big)$ is a face of $C(d,n)$.*

Proof. We will find a non-zero vector $c = (\gamma_1, \ldots, \gamma_d)$ and a number α such that

$$\langle c, v_i \rangle = \alpha \quad \text{for} \quad i \in I \qquad \text{and} \qquad \langle c, v_i \rangle < \alpha \quad \text{for} \quad i \notin I.$$

This would clearly imply that F is a face of $C(d,n)$.

Let $I = \{i_1, i_2, \ldots, i_k\}$, hence $k \leq d/2$. Let us consider a polynomial p of degree d in τ:

$$p(\tau) = -\tau^{d-2k}(\tau - \tau_{i_1})^2 \cdots (\tau - \tau_{i_k})^2.$$

Hence $p(\tau_i) = 0$ for $i \in I$ and $p(\tau_i) < 0$ for $i \in \{1, \ldots, n\} \setminus I$. Let us write

$$p(\tau) = \gamma_d \tau^d + \gamma_{d-1}\tau^{d-1} + \ldots + \gamma_1 \tau_1 - \alpha$$

and let $c = (\gamma_1, \ldots, \gamma_d)$. Then

$$\langle c, q(\tau) \rangle = \gamma_1 \tau + \gamma_2 \tau^2 + \ldots + \gamma_d \tau^d = p(\tau) + \alpha.$$

Hence $\langle c, v_i \rangle = \alpha$ for $i \in I$ and $\langle c, v_i \rangle < \alpha$ for $i \notin I$ and the proof follows. $\square$

PROBLEMS.

1. Let $P = \mathrm{conv}(v_1, \ldots, v_n) \subset \mathbb{R}^d$ be a d-dimensional polytope with n vertices and let $k > d/2$. Prove that if every k vertices $v_{i_1}, \ldots, v_{i_k}$ are the vertices of a face of P, then $n = d + 1$.

Hint: Use Radon's Theorem (see Theorem 3.1).

2. Let $I \subset \{1, \ldots, n\}$ be a set, $|I| = d$. Prove that $F = \mathrm{conv}(v_i : i \in I)$ is a facet of $C(d, n)$ if and only if the following *Gale's evenness condition* is satisfied: every two elements i, j which are not in I are separated by an even number of elements from I.

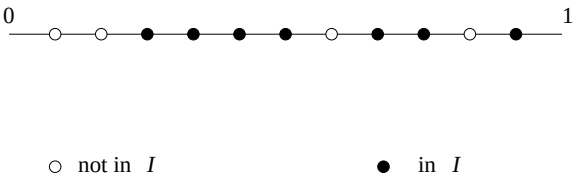


Figure 50

3. Describe the faces of the polytope $C(4, n)$.

4*. Prove that for $1 \le k \le d-1$, the number f_k of k-dimensional faces of $C(d, n)$ is

$$f_k = \begin{cases} \sum_{j=1}^{(d+1)/2} \frac{n}{n-j} \binom{n-j}{j} \binom{j}{k+1-j} & \text{if } d \text{ is odd} \\ \sum_{j=0}^{d/2} \frac{k+2}{n-j} \binom{n-j}{j+1} \binom{j+1}{k+1-j} & \text{if } d \text{ is even.} \end{cases}$$

5. Let us define a closed curve

$$\phi(\tau) = \Big(\sin\tau, \cos\tau, \sin 2\tau, \cos 2\tau, \ldots, \sin k\tau, \cos k\tau\Big), \quad 0 \le \tau \le 2\pi$$

in $\mathbb{R}^d$, $d = 2k$. Let $0 < \tau_1 < \tau_2 < \ldots < \tau_n$ and let $v_i = \phi(\tau_i)$. Let $P = \mathrm{conv}(v_1, \ldots, v_n)$. Prove that P and $C(d, n)$ have the same facial structure. In particular, the facial structure of P does not depend on the choice of $0 < \tau_1 < \ldots < \tau_n < 2\pi$ and the cyclic permutation $v_1 \longmapsto v_2 \longmapsto \ldots \longmapsto v_n \longmapsto v_1$ gives rise to a permutation of faces of $C(d, n)$.

The cyclic polytope $C(d, n)$ has the following important property: among all polytopes in $\mathbb{R}^d$ with n vertices, the polytope $C(d, n)$ has the largest number of faces of each dimension $k = 1, \ldots, d-1$, see P. McMullen and G.C. Shephard, *Convex Polytopes and the Upper Bound Conjecture*, Cambridge University Press, 1971. Our next goal is to sketch a proof of this result.

38. Simple Polytopes

In this section, we introduce an important class of polytopes.

(38.1) Definitions. A d-dimensional polytope P is called *simple* if every vertex v of P belongs to exactly d facets of P. In particular, if a polytope $P \subset \mathbb{R}^d$, $\dim P = d$, is defined as a d-dimensional bounded polyhedron,

$$P = \big\{x \in \mathbb{R}^d : \langle c_i, x\rangle \le \beta_i : i = 1, \ldots, m\big\},$$

then P is simple provided for every vertex v of P, the set $I(v) = \{i : \langle c_i, v\rangle = \beta_i\}$ of inequalities that are active on v consists of precisely d elements (cf. Theorem 11.2).

Let v, u be vertices of a polytope P (not necessarily simple). If $[u, v]$ is an edge of P, we say that u and v are *neighbors*.

PROBLEMS.

1°. Prove that the cube $I = \{x \in \mathbb{R}^d : -1 \le \xi_i \le 1 \text{ for } i = 1, \dots, d\}$ is a simple polytope and that the (hyper)octahedron $O = \{x \in \mathbb{R}^d : |\xi_1| + \dots + |\xi_d| \le 1\}$ is not simple for $d > 2$.

2. A polytope which is the the convex hull of affinely independent points is called a *simplex*. Let $P \subset \mathbb{R}^d$ be a polytope such that $0 \in \operatorname{int} P$. Prove that P is simple if and only if every facet of P° is a simplex (such polytopes are called simplicial).

3. Prove that if a polytope is both simple and simplicial (see Problem 2 above) then it is either a simplex or a polygon.

4°. Let us fix a vector $a = (\alpha_1, \dots, \alpha_d)$ with distinct coordinates and let $P(a)$ be the corresponding permutation polytope (see Section 35). Prove that $P(a)$ is a simple $(d-1)$-dimensional polytope.

5. Prove that the cyclic polytope $C(d, n)$ is simplicial.

6. Let $P \subset \mathbb{R}^d$ be a polytope, let v be a vertex of P and let $u_1, \dots, u_n$ be the neighbors of P. Suppose that $\ell : \mathbb{R}^d \longrightarrow \mathbb{R}$ is a linear function such that $\ell(v) > \ell(u_i)$ for $i = 1, \dots, n$. Prove that the maximum of ℓ on P is attained at v.

Locally, in the neighborhood of every vertex, a simple d-dimensional polytope P looks like the standard non-negative orthant

$$\mathbb{R}^d_+ = \{(\xi_1, \dots, \xi_d) : \xi_i \ge 0, i = 1, \dots, d\}.$$

(38.2) Proposition. *Let $P \subset \mathbb{R}^d$ be a d-dimensional simple polytope. Then*

(1) *Every vertex v of P has precisely d neighbors $u_1, \dots, u_d$;*
(2) *For every vertex v of P and for every $k < d$ neighbors $u_{i_1}, \dots, u_{i_k}$ of v there exist a unique k-dimensional face F of P containing v and $u_{i_1}, \dots, u_{i_k}$;*
(3) *The intersection of any $k \le d$ facets of P containing v is a $(d-k)$ dimensional face of P;*
(4) *Let $\ell : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function such that $\ell(u_i) < \ell(v)$ for all neighbors u_i of some vertex v of P. Then the maximum of ℓ on P is attained at the point v.*
(5) *Every face of P is a simple polytope.*

Proof. Let $P = \{x \in \mathbb{R}^d : \langle c_i, x\rangle \le \beta_i, i = 1, \dots, m\}$ and let $v \in P$ be a vertex. Translating P, if necessary, we may assume that $v = 0$ is the origin, and changing the basis, if necessary, we may assume that the inequalities that are active on v are $\xi_i \ge 0$ for $i = 1, \dots, d$. Thus P is the intersection of the non-negative orthant

$\mathbb{R}^d_+ = \{x : \xi_i \geq 0,\ i = 1, \ldots, d\}$ with finitely many halfspaces containing the origin in their interior.

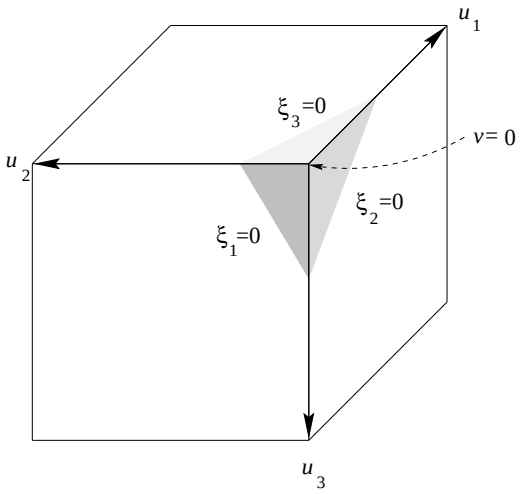


Figure 51

To prove (1), note that v belongs to exactly d edges; the i-th edge is the intersection of P with the hyperplane $\sum_{j \neq i} \xi_j = 0$. The other endpoints of these edges are the neighbors $u_1, \ldots, u_d$ of v. Thus all but the i-th coordinate of u_i are 0 and the i-th coordinate of u_i is positive. Let us prove that there are no other edges with the endpoint $v = 0$. Indeed, if $c = (\gamma_1, \ldots, \gamma_d)$ is a vector such that $0 = \langle c, v\rangle \geq \langle c, x\rangle$ for all $x \in P$, so that $F = \{x : \langle c, x\rangle = 0\}$ is a face of P containing v, substituting $x = u_i$, we get that $\gamma_i \leq 0$ for $i = 1, \ldots, d$. Then $F = \{x \in P : \sum_{i:\gamma_i<0} \xi_i = 0\}$, which is an edge of P if and only if all but one γ_i are equal to 0.

To prove (2), note that F is the intersection of P with the hyperplane $\sum_{j \neq i_1, \ldots, i_k} \xi_j = 0$.

To prove (3), we observe that the i-th facet is defined by the equation $\xi_i = 0$. Therefore, the intersection of the facets with the indices i_1, $i_2, \ldots, i_k$ is the face F defined by the equation $\sum_{j=i_1,\ldots,i_k} \xi_j = 0$. Clearly $\dim F = k$, since F contains a point, where $\xi_j > 0$ for $j \neq i_1, \ldots, i_k$.

To prove (4), assume that $\ell(x) = \gamma_1\xi_1 + \ldots + \gamma_d\xi_d$. Hence we have $\ell(u_i) < \ell(v) = 0$, so all γ_i's are negative. Therefore $\ell(x) \leq 0$ for any $x \in \mathbb{R}^d_+$ and $\ell(x) = 0$ if and only if $x = 0$.

To prove (5), it suffices to prove that if F is a facet of P then F is a $(d-1)$-dimensional simple polytope. Let us choose a vertex v of F. Then v is a vertex of P and there are exactly d facets $G_1, \ldots, G_d$ of P containing v. Without loss of generality, we may assume that $F = G_d$. As above, we may assume that $v = 0$ and that G_i is defined by the equation $\xi_i = 0$. Then $F \cap G_i$, $i = 1, \ldots, d-1$ are the facets of F containing v (cf. Part 3). □

PROBLEMS.

1*. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope with n facets. Construct a simple

d-dimensional polytope $\tilde{P} \subset \mathbb{R}^d$ with n facets such that $f_i(P) \le f_i(\tilde{P})$ for $i = 0, \dots, d-2$, where f_i is the number of i-dimensional faces.

Hint: Suppose that

$$P = \big\{x : \langle c_i, x\rangle \le \beta_i,\ i = 1, \dots, n\big\}.$$

Let

$$\tilde{P} = \big\{x : \langle c_i, x\rangle \le \tilde{\beta}_i,\ i = 1, \dots, n\big\},$$

where $\tilde{\beta}_i$ are generic small perturbations of β_i.

2°. Prove that every face of a simplicial polytope is a simplex (see Problem 2, Section 38.1).

3. Prove that a d-dimensional polytope is simple if every vertex of the polytope has precisely d neighbors.

4. Let P be a simple d-dimensional polytope. Prove that $df_0(P) = 2f_1(P)$.

5. Let P be a 3-dimensional simple polytope. Then the facets of P are polygons and let p_k be the number of k-gons among its facets. Prove that $3p_3 + 2p_4 + p_5 = 12 + \sum_{k\ge 7}(k-6)p_k$.

6*. Let $p_k : k > 2, k \ne 6$ be non-negative integers such that $3p_3 + 2p_4 + p_5 = 12 + \sum_{k\ge 7}(k-6)p_k$. Prove that for some p_6 there exists a 3-dimensional polytope whose facets consist of p_k k-gons for $k = 3, 4, 5, \dots$.

39. The h-vector of a Simple Polytope. Dehn-Sommerville Equations

(39.1) Definition. Let P be a a d-dimensional simple polytope and let $f_i(P)$ be the number of the i-dimensional faces of P (we agree that $f_d(P) = 1$). Let

$$h_k(P) = \sum_{i=k}^{d} (-1)^{i-k} \binom{i}{k} f_i(P) \quad \text{for} \quad k = 0, \dots, d,$$

where $\binom{m}{n} = \dfrac{m!}{n!(m-n)!}$ is the binomial coefficient.

The $(d+1)$-tuple $\big(h_0(P), \dots, h_d(P)\big)$ is called the *h-vector* of P.

(39.2) Lemma. *Let P be a d-dimensional simple polytope. Then*

$$f_i(P) = \sum_{k=i}^{d} \binom{k}{i} h_k(P) \quad \textit{for} \quad i = 0, \dots, d.$$

Furthermore, the numbers $h_k(P)$ are uniquely determined by the above equations.

Proof. Let us introduce two polynomials in one variable τ:

$$\mathbf{f}(\tau) = \sum_{i=0}^{d} f_i(P)\tau^i \quad \text{and} \quad \mathbf{h}(\tau) = \sum_{k=0}^{d} h_k(P)\tau^k.$$

Then the equations of Definition 39.1 are equivalent to the identity $\mathbf{f}(\tau - 1) = \mathbf{h}(\tau)$, whereas the equations of Lemma 39.2 are equivalent to the identity $\mathbf{f}(\tau) = \mathbf{h}(\tau + 1)$. The proof now follows. □

Now is the main result about the h-vector.

(39.3) Theorem. *Let $P \subset \mathbb{R}^d$ be a simple d-dimensional polytope. Let $\ell : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function such that $\ell(v_i) \neq \ell(v_j)$ for every pair of neighbors v_i, v_j. For a vertex v of P let us define the index of v with respect to ℓ as the number of neighbors v_i of v such that $\ell(v_i) < \ell(v)$. Then for every $0 \leq k \leq d$ the number of vertices of P with the index k with respect to ℓ is equal to $h_k(P)$. In particular, this number does not depend on the function ℓ.*

Proof. Let us count the i-dimensional faces F of P. On one hand, we know that this number is $f_i(P)$. On the other hand, on every i-dimensional face F let us mark its only vertex $v \in F$ where the maximum of ℓ on F is attained: $\ell(v) = \max_{x \in F} \ell(x)$ (cf. Part 4 of Proposition 38.2). Let us see how many times a given vertex $v \in P$ gets marked.

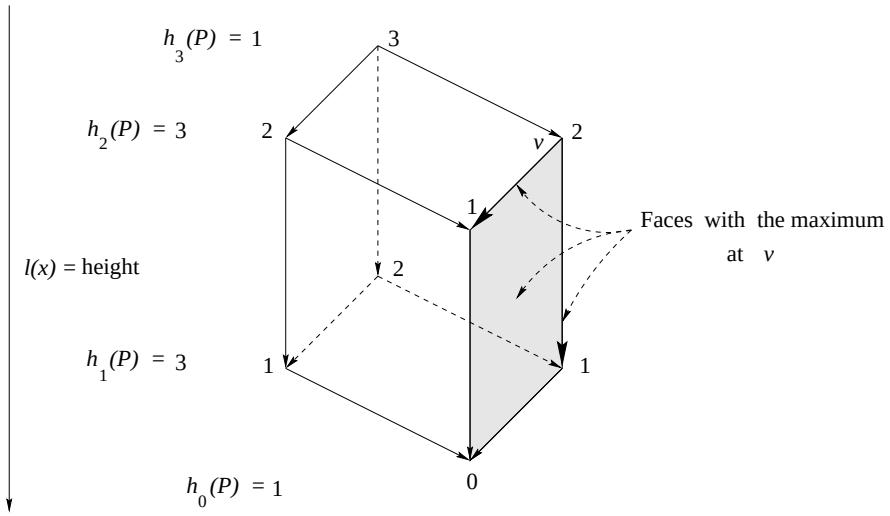


Figure 52

Suppose that the index of v is equal to $k \geq i$. Then there are exactly k neighbors v_j of v for which $\ell(v_j) < \ell(v)$. By Part 2 of Proposition 38.2, for any i of these neighbors there exists a unique i-dimensional face F containing them and by Part 4 of Proposition 38.2, ℓ attains its maximum on F at v. Conversely, if F is an i-dimensional face such that the maximum of ℓ on F is attained at v, then v should have i neighbors v_j (see Parts 5 and 1 of Proposition 38.2) and for every such a neighbor v_j we must have $\ell(v_j) < \ell(v)$. In particular, the index of v with respect to ℓ must be at least i.

Summarizing, we observe that every vertex v of index $k \geq i$ is marked precisely $\binom{k}{i}$ times and that the vertices whose indices are smaller than i are not marked at all. Denoting for a moment by $h_k(P, \ell)$ the number of vertices of P whose index with respect to ℓ is k, we conclude that

$$f_i(P) = \sum_{k=i}^{d} \binom{k}{i} h_k(P, \ell) \quad \text{for} \quad i = 0, \ldots, d.$$

By Lemma 39.2, the numbers $h_k(P, \ell)$ are uniquely determined by the above identities and must coincide with $h_k(P)$. $\square$

(39.4) Corollary (Dehn - Sommerville Equations). *Let P be a simple d-dimensional polytope. Then*

$$h_k(P) = h_{d-k}(P) \quad \textit{for} \quad k = 0, \ldots, d.$$

Proof. Let us choose a linear function $\ell : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that $\ell(v_i) \neq \ell(v_j)$ for any two neighboring vertices v_i, v_j of P. Then $h_k(P)$ is the number of vertices of P that have index k with respect to ℓ (Theorem 39.3). Applying Theorem 39.3 with the function $-\ell$, we conclude that the number of vertices of P with the index $d-k$ with respect to $-\ell$ is equal to $h_{d-k}(P)$. However, the index of a vertex with respect to ℓ is k if and only if the index of the vertex with respect to $-\ell$ is $d-k$. The proof now follows. $\square$

The relations (in a different form) were found by M. Dehn in 1905 for $d \leq 5$ and for an arbitrary d by D. Sommerville in 1927. The proof belongs to P. McMullen (1970).

PROBLEMS.

1°. Prove that $\sum_{k=0}^{d} h_k(P) = f_0(P)$.

2°. Let P be a simple d-dimensional polytope. Check that the equation $h_0(P) = h_d(P)$ is the Euler-Poincaré formula for P.

3°. Prove that $h_k(P) \geq 0$ for every simple polytope P.

4. Compute the h-vector of a d-dimensional simplex (the convex hull of $d+1$ affinely independent points in $\mathbb{R}^d$).

5. Let P be a simple 3-dimensional polytope. Prove that the Dehn-Sommerville equations for P are equivalent to $f_0 - f_1 + f_2 = 2$ and $3f_0 - 2f_1 = 0$. In particular, the number of faces of a simple 3-dimensional polytope is determined by the number of vertices.

6. Let P be a simple 4-dimensional polytope. Prove that the Dehn-Sommerville equations for P are equivalent to $f_0 - f_1 + f_2 - f_3 = 0$ and $f_1 = 2f_0$.

7. Let $I = \{(\xi_1, \ldots, \xi_d) : 0 \leq \xi_i \leq 1\}$ be a d-dimensional cube. Prove that $h_k(I) = \binom{d}{k}$ for $k = 0, \ldots, d$.

8. For a permutation $\sigma \in S_n$, let us define it *descend* as a number $i = 2, \ldots, n$ such that $\sigma(i) < \sigma(i-1)$. Let $E(n,k)$ be the number of permutations having precisely $k-1$ descends, $k = 1, \ldots, n$. Let $a = (\alpha_1, \ldots, \alpha_n)$ be a point with distinct coordinates $\alpha_1 > \ldots > \alpha_n$ and let $P = P(a)$ be the corresponding permutation polytope (see Section 35). Then $P(a)$ is a simple $(n-1)$-dimensional polytope (see Problem 4, Section 38.1) and one can define $h_h(P)$ for $k = 0, \ldots, n-1$. Prove that $h_k(P) = E(n, k+1)$ and that $E(n,k) = E(n, n-k+1)$ for $k = 1, \ldots, n$. The numbers $E(n,k)$ are called *Eulerian numbers*.

40. The Upper Bound Theorem

In this section, we sketch a proof of the Upper Bound Theorem, which tells us the maximum number of faces that a polytope of a given dimension and with a given number of vertices (equivalently, facets) may have. The proof below belongs to P. McMullen (1970). More precisely, McMullen gave a detailed proof which is, in a sense, dual to our approach, but he also remarked what the dual (to his proof, that is, our proof) looks like. In detail, our approach is described in A. Brønsted, *An Introduction to Convex Polytopes*, Springer-Verlag, 1983.

(40.1) Lemma. *Let P be a d-dimensional simple polytope and let F be its facet. Then F is a $(d-1)$-dimensional simple polytope and*

$$h_k(F) \leq h_k(P) \quad \textit{for} \quad k = 0, \ldots, d-1.$$

Moreover, if the intersection of every $k+1$ facets of P is non-empty, then $h_k(F) = h_k(P)$.

Proof. Part 5 of Proposition 38.2 implies that F is a simple polytope. Since F is a facet of P, there is a linear function ℓ and a number α such that $\ell(x) \geq \alpha$ for any $x \in P$ and $\ell(x) = \alpha$ if and only if $x \in F$. In particular, $\ell(v) > \alpha$ for any vertex of v of P that is not in F. Let us perturb $\ell \longmapsto \tilde{\ell}$ slightly so that $\tilde{\ell}$ has different values on different vertices of P and $\tilde{\ell}(v) > \tilde{\ell}(u)$ for every pair of vertices u and v of P, where $v \notin F$ and $u \in F$.

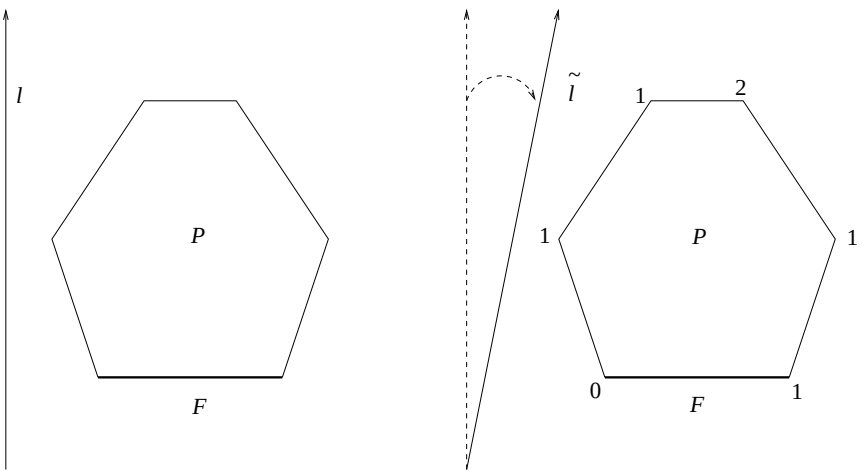


Figure 53

Theorem 39.3 asserts that $h_k(P)$ (resp. $h_k(F)$) is the number of vertices of P (resp. of F) that have index k with respect to $\tilde{\ell}$. Let v be a vertex of F of index k. Then v has $(d-1)$ neighbors in $u_1, \ldots, u_{d-1}$ in F and precisely k of them have a smaller value of $\tilde{\ell}$. As a vertex of P, v has one additional neighbor $u_d \notin F$, where we have $\tilde{\ell}(u_d) > \tilde{\ell}(v)$. Therefore, the index of v in P is also k.

Suppose that the intersection of every k facets of P is non-empty. Suppose that v is a vertex of P of index k and v is *not* in F. There are $d-k$ neighbors, say, $u_{k+1}, \ldots, u_d$ of v such that $\tilde{\ell}(u_i) > \tilde{\ell}(v)$ for all $i = k+1, \ldots, d$. Then, by Part 2 of Proposition 38.2, there is $(d-k)$-dimensional face G containing v and $u_{k+1}, \ldots, u_d$ and by Part 4 it follows that $\tilde{\ell}$ achieves its minimum value on G at v. Moreover, using Part 3 of Proposition 38.2, we can represent G as the intersection of k facets, $G = F_1 \cap \ldots \cap F_k$, where F_i is a facet of P containing v and all but u_i of its neighbors. Then the intersection $F \cap F_1 \cap \ldots \cap F_k = F \cap G$ is non-empty, and hence we must have $\tilde{\ell}(x) > \tilde{\ell}(v)$ for some $x \in F$, which is a contradiction.

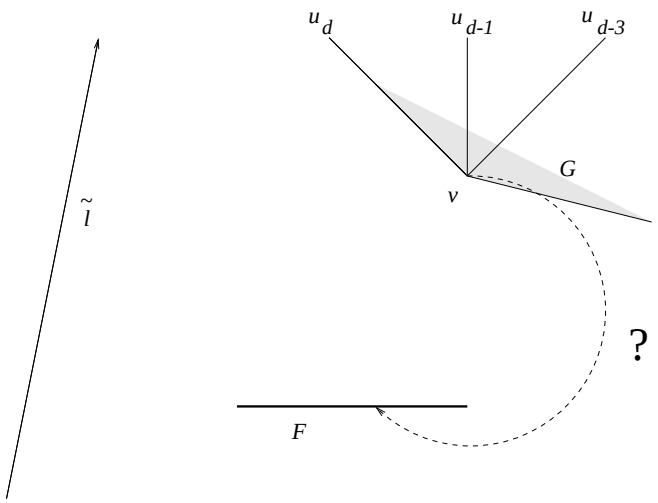


Figure 54

The contradiction shows that there are no points of index k outside of F and hence $h_k(F) = h_k(P)$. □

(40.2) Lemma. *Let P be a simple d-dimensional polytope. Then*

$$\sum_{F \text{ is a facet of } P} h_k(F) = (d-k)h_k(P) + (k+1)h_{k+1}(P).$$

Proof. To prove the formula, let us choose a linear function ℓ, which attains different values at neighbors in P. By Theorem 39.3, the left hand side of the formula counts the vertices on the facets of P that have index k relative to the facet. Let us us see how many times a given vertex v of P gets counted. There are exactly d facets of P containing v and each such a facet F of P contains all but one neighbor of v (see Part 2 of Proposition 38.2).

If the index of v in P with respect to ℓ is smaller than k then there are fewer than k neighbors u_i of v with $\ell(u_i) < \ell(v)$, so v is not counted at all in the left hand side.

Similarly, if the index of v is greater than $k+1$, then every facet F contains at least $k+1$ neighbors u_i of v with $\ell(u_i) < \ell(v)$, and so v is not counted in the left hand side.

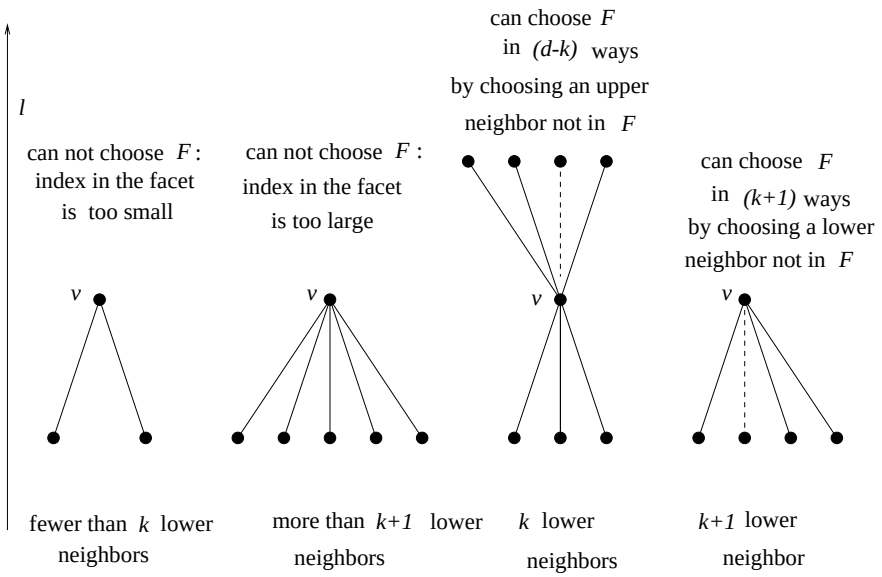


Figure 55

If the index of v in P is k then there there are precisely k neighbors $u_1, \ldots, u_k$ of v with $\ell(u_i) < \ell(v)$. Then there are precisely $d-k$ facets F of P that contain v and all of $u_1, \ldots, u_k$. Hence every vertex v of P of the index k is counted $(d-k)$ times.

Finally, suppose that the index of v in P is $k+1$. Then there are exactly $k+1$ facets F of P such that the index of v relative to the facet is k. The i-th such facet F is determined uniquely by the condition that it contains v and all neighbors of v except one neighbor u_i such that $\ell(u_i) < \ell(v)$ is smaller. Hence a vertex v of P of index $k+1$ is counted $(k+1)$ times. The proof now follows. □

(40.3) Theorem. *Let P be a d-dimensional simple polytope with n facets. Then*

$$h_k(P) \leq \binom{n-d+k-1}{k} \quad \textit{for} \quad k = 0, \ldots, d.$$

Moreover, if the intersection of every k facets of P is non-empty, then

$$h_k(P) = \binom{n-d+k-1}{k}.$$

Proof. Combining Lemma 40.1 and Lemma 40.2, we get we get

$$nh_i(P) \geq (d-i)h_i(P) + (i+1)h_{i+1}(P)$$

or

$$h_{i+1}(P) \le \frac{n-d+i}{i+1} h_i(P) \quad \text{for} \quad k = 0, \dots, d,$$

and the equality holds if every $i+1$ facets of P intersect.

Combining it with $h_0(P) = h_d(P) = 1$, we get the desired inequality. □

(40.4) The Upper Bound Theorem. *Let P be a d-dimensional polytope with n facets and let Q be a d-dimensional simple polytope with n facets, such that for every $k \le d/2$, the intersection of k facets of Q is a $(d-k)$-dimensional face of Q. For example, one can choose Q to be the polar of the cyclic polytope $C(d,n)$ with the origin chosen in the interior of $C(d,n)$. Then*

$$f_i(P) \le f_i(Q) \quad \textit{for all} \quad i = 0, \dots, d-1.$$

Besides, for $i \le \lfloor d/2 \rfloor$,

$$f_i(Q) = \sum_{k=i}^{\lfloor d/2 \rfloor} \binom{k}{i} \binom{n-d+k-1}{k} + \sum_{k=\lfloor d/2 \rfloor + 1}^{d} \binom{k}{i} \binom{n-k-1}{d-k} \tag{40.4.1}$$

and for $i > \lfloor d/2 \rfloor$,

$$f_i(Q) = \binom{n}{d-i}. \tag{40.4.2}$$

Proof. By Proposition 37.2, every $k \le d/2$ vertices of $C(d,n)$ belong to a proper face of $C(d,n)$ of dimension $k-1$. Theorem 34.3 implies that for every $k \le d/2$, the intersection k facets of $Q = C^\circ(d,n)$ is a $(d-k)$-dimensional face of Q.

Then (40.4.2) is obvious. By Lemma 40.3,

$$h_k(Q) = \binom{n-d+k-1}{k} \quad \text{for} \quad k = 0, \dots, \lfloor d/2 \rfloor. \tag{40.4.3}$$

Combining this with the Dehn-Sommerville equations $h_k(Q) = h_{d-k}(Q)$ (Corollary 39.4) for $k > d/2$, we get (40.4.1) from Lemma 39.2.

If P is a simple polytope, then from Theorem 40.3 and (40.4.3), we have $h_k(P) \le h_k(Q)$ for $k \le d/2$. Combining this with the Dehn-Sommerville equations, we get $h_k(P) \le h_k(Q)$ for all $k = 0, \dots, d$. Then by Lemma 39.2, $f_i(P) \le f_i(Q)$ for all $i = 1, \dots, d-1$.

The case of an arbitrary polytope P reduces to that of a simple polytope P by Problem 1, Section 38.2. □

PROBLEMS.

1. Let us fix the dimension d. Prove that the number of faces of a polytope $P \subset \mathbb{R}^d$ with n facets is $O(n^{d/2})$.

2. Deduce the Upper Bound Theorem for d-dimensional polytopes with n vertices.

3. Write the Upper Bound inequalities for a 4-dimensional polytope with n facets and for a 4-dimensional polytope with n vertices.

4^*. Let P be a simple d-dimensional polytope. Prove the Unimodality Theorem: $h_0(P) \leq h_1(P) \leq h_2(P) \leq \ldots \leq h_{\lfloor d/2 \rfloor}(P)$.