

Jiří Matoušek

Lectures on Discrete Geometry

Excerpt

Springer-Verlag

Berlin Heidelberg New York

London Paris Tokyo

Hong Kong Barcelona

Budapest

Contents

Notation and Terminology	iii
1 Convexity	1
1.1 Linear and Affine Subspaces, General Position	1
1.2 Convex Sets, Convex Combinations, Separation	5
1.3 Radon's Lemma and Helly's Theorem	9
1.4 Centerpoint and Ham Sandwich	14
5 Convex Polytopes	17
5.1 Geometric Duality	18
5.2 H -Polytopes and V -Polytopes	22
5.3 Faces of a Convex Polytope	26
5.4 Many Faces: The Cyclic Polytopes	32
5.5 The Upper Bound Theorem	35
5.7 Voronoi Diagrams	40
6 Number of Faces in Arrangements	51
6.1 Arrangements of Hyperplanes	52
6.2 Arrangements of Other Geometric Objects	55
Bibliography	67

Notation and Terminology

This section summarizes rather standard things, and it is mainly for reference. More special notions are introduced gradually throughout the book. In order to facilitate independent reading of various parts, some of the definitions are even repeated several times.

If X is a set, $|X|$ denotes the number of elements (cardinality) of X . If X is a *multiset*, in which some elements may be repeated, then $|X|$ counts each element with its multiplicity.

The very slowly growing function $\log^* x$ is defined by $\log^* x = 0$ for $x \leq 1$ and $\log^* x = 1 + \log^*(\log_2 x)$ for $x > 1$.

For a real number x , $\lfloor x \rfloor$ denotes the largest integer less than or equal to x , and $\lceil x \rceil$ means the smallest integer greater than or equal to x . The boldface letters \mathbf{R} and \mathbf{Z} stand for the real numbers and for the integers, respectively, while \mathbf{R}^d denotes the d -dimensional Euclidean space. For a point $x = (x_1, x_2, \dots, x_d) \in \mathbf{R}^d$, $\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$ is the Euclidean norm of x , and for $x, y \in \mathbf{R}^d$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_d y_d$ is the scalar product. Points of \mathbf{R}^d are usually considered as column vectors.

The symbol $B(x, r)$ denotes the closed ball of radius r centered at x in some metric space (usually in \mathbf{R}^d with the Euclidean distance), i.e., the set of all points with distance at most r from x . We write B^n for the unit ball $B(0, 1)$ in \mathbf{R}^n . The symbol ∂A denotes the boundary of a set $A \subseteq \mathbf{R}^d$, that is, the set of points at zero distance from both A and its complement.

For a measurable set $A \subseteq \mathbf{R}^d$, $\text{vol}(A)$ is the d -dimensional Lebesgue measure of A (in most cases the usual volume).

Let f and g be real functions (of one or several variables). The notation $f = O(g)$ means that there exists a number C such that $|f| \leq C|g|$ for all values of the variables. Normally, C should be an absolute constant, but if f and g depend on some parameter(s) that we explicitly declare to be fixed (such as the space dimension d), then C may depend on these parameters as well. The notation $f = \Omega(g)$ is equivalent to $g = O(f)$, $f(n) = o(g(n))$ to $\lim_{n \rightarrow \infty} (f(n)/g(n)) = 0$, and $f = \Theta(g)$ means that both $f = O(g)$ and $f = \Omega(g)$.

For a random variable X , the symbol $\mathbf{E}[X]$ denotes the expectation of X , and $\text{Prob}[A]$ stands for the probability of an event A .

Graphs are considered simple and undirected in this book unless stated otherwise, so a graph G is a pair (V, E) , where V is a set (the *vertex set*) and $E \subseteq \binom{V}{2}$ is the *edge set*. Here $\binom{V}{k}$ denotes the set of all k -element subsets of V . For a *multigraph*, the edges form a multiset, so two vertices can be connected by several edges. For a given (multi)graph G , we write $V(G)$ for the vertex set and $E(G)$ for the edge set. A *complete graph* has all possible edges; that is, it is of the form $(V, \binom{V}{2})$. A complete graph on n vertices is denoted by K_n . A graph G is *bipartite* if the vertex set can be partitioned into two subsets V_1 and V_2 , the (*color*) *classes*, in such a way that each edge connects a vertex of V_1 to a vertex of V_2 . A graph $G' = (V', E')$ is a *subgraph* of a graph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. We also say that G *contains a copy* of H if there is a subgraph G' of G isomorphic to H , where G' and H are *isomorphic* if there is a bijective map $\varphi: V(G') \rightarrow V(H)$ such that $\{u, v\} \in E(G')$ if and only if $\{\varphi(u), \varphi(v)\} \in E(H)$ for all $u, v \in V(G')$. The *degree* of a vertex v in a graph G is the number of edges of G containing v . An *r -regular graph* has all degrees equal to r . Paths and cycles are graphs as in the following picture,



and a path or cycle in G is a subgraph isomorphic to a path or cycle, respectively. A graph G is *connected* if every two vertices can be connected by a path in G .

We recall that a set $X \subseteq \mathbf{R}^d$ is *compact* if and only if it is closed and bounded, and that a continuous function $f: X \rightarrow \mathbf{R}$ defined on a compact X attains its minimum (there exists $x_0 \in X$ with $f(x_0) \leq f(x)$ for all $x \in X$).

The *Cauchy–Schwarz inequality* is perhaps best remembered in the form $\langle x, y \rangle \leq \|x\| \cdot \|y\|$ for all $x, y \in \mathbf{R}^n$.

A real function f defined on an interval $A \subseteq \mathbf{R}$ (or, more generally, on a convex set $A \subseteq \mathbf{R}^d$) is *convex* if $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ for all $x, y \in A$ and $t \in [0, 1]$. Geometrically, the graph of f on $[x, y]$ lies below the segment connecting the points $(x, f(x))$ and $(y, f(y))$. If the second derivative satisfies $f''(x) \geq 0$ for all x in an (open) interval $A \subseteq \mathbf{R}$, then f is convex on A . *Jensen's inequality* is a straightforward generalization of the definition of convexity: $f(t_1x_1 + t_2x_2 + \cdots + t_nx_n) \leq t_1f(x_1) + t_2f(x_2) + \cdots + t_nf(x_n)$ for all choices of nonnegative t_i summing to 1 and all $x_1, \dots, x_n \in A$. Or in integral form, if μ is a probability measure on A and f is convex on A , we have $f(\int_A x d\mu(x)) \leq \int_A f(x) d\mu(x)$. In the language of probability theory, if X is a real random variable and $f: \mathbf{R} \rightarrow \mathbf{R}$ is convex, then $f(\mathbf{E}[X]) \leq \mathbf{E}[f(X)]$; for example, $(\mathbf{E}[X])^2 \leq \mathbf{E}[X^2]$.

1

Convexity

We begin with a review of basic geometric notions such as hyperplanes and affine subspaces in \mathbf{R}^d , and we spend some time by discussing the notion of general position. Then we consider fundamental properties of convex sets in \mathbf{R}^d , such as a theorem about the separation of disjoint convex sets by a hyperplane and Helly's theorem.

1.1 Linear and Affine Subspaces, General Position

Linear subspaces. Let \mathbf{R}^d denote the d -dimensional Euclidean space. The points are d -tuples of real numbers, $x = (x_1, x_2, \dots, x_d)$.

The space \mathbf{R}^d is a vector space, and so we may speak of linear subspaces, linear dependence of points, linear span of a set, and so on. A linear subspace of \mathbf{R}^d is a subset closed under addition of vectors and under multiplication by real numbers. What is the geometric meaning? For instance, the linear subspaces of \mathbf{R}^2 are the origin itself, all lines passing through the origin, and the whole of \mathbf{R}^2 . In \mathbf{R}^3 , we have the origin, all lines and planes passing through the origin, and \mathbf{R}^3 .

Affine notions. An arbitrary line in \mathbf{R}^2 , say, is *not* a linear subspace unless it passes through 0. General lines are what are called *affine subspaces*. An affine subspace of \mathbf{R}^d has the form $x + L$, where $x \in \mathbf{R}^d$ is some vector and L is a linear subspace of \mathbf{R}^d . Having defined affine subspaces, the other “affine” notions can be constructed by imitating the “linear” notions.

What is the *affine hull* of a set $X \subseteq \mathbf{R}^d$? It is the intersection of all affine subspaces of \mathbf{R}^d containing X . As is well known, the linear span of a set X can be described as the set of all linear combinations of points of X . What is an *affine combination* of points $a_1, a_2, \dots, a_n \in \mathbf{R}^d$ that would play an analogous role? To see this, we translate the whole set by $-a_n$, so that a_n becomes the origin, we make a linear combination, and we translate back by

$+a_n$. This yields an expression of the form $\beta_1(a_1 - a_n) + \beta_2(a_2 - a_n) + \cdots + \beta_n(a_n - a_n) + a_n = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_{n-1} a_{n-1} + (1 - \beta_1 - \beta_2 - \cdots - \beta_{n-1}) a_n$, where β_1, \dots, β_n are arbitrary real numbers. Thus, an affine combination of points $a_1, \dots, a_n \in \mathbf{R}^d$ is an expression of the form

$$\alpha_1 a_1 + \cdots + \alpha_n a_n, \text{ where } \alpha_1, \dots, \alpha_n \in \mathbf{R} \text{ and } \alpha_1 + \cdots + \alpha_n = 1.$$

Then indeed, it is not hard to check that the affine hull of X is the set of all affine combinations of points of X .

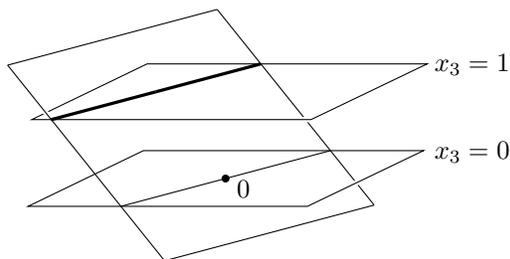
The *affine dependence* of points a_1, \dots, a_n means that one of them can be written as an affine combination of the others. This is the same as the existence of real numbers $\alpha_1, \alpha_2, \dots, \alpha_n$, at least one of them nonzero, such that both

$$\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n = 0 \text{ and } \alpha_1 + \alpha_2 + \cdots + \alpha_n = 0.$$

(Note the difference: In an affine *combination*, the α_i sum to 1, while in an affine *dependence*, they sum to 0.)

Affine dependence of a_1, \dots, a_n is equivalent to linear dependence of the $n-1$ vectors $a_1 - a_n, a_2 - a_n, \dots, a_{n-1} - a_n$. Therefore, the maximum possible number of affinely independent points in \mathbf{R}^d is $d+1$.

Another way of expressing affine dependence uses “lifting” one dimension higher. Let $b_i = (a_i, 1)$ be the vector in \mathbf{R}^{d+1} obtained by appending a new coordinate equal to 1 to a_i ; then a_1, \dots, a_n are affinely dependent if and only if b_1, \dots, b_n are linearly dependent. This correspondence of affine notions in \mathbf{R}^d with linear notions in \mathbf{R}^{d+1} is quite general. For example, if we identify \mathbf{R}^2 with the plane $x_3 = 1$ in \mathbf{R}^3 as in the picture,



then we obtain a bijective correspondence of the k -dimensional linear subspaces of \mathbf{R}^3 that do not lie in the plane $x_3 = 0$ with $(k-1)$ -dimensional affine subspaces of \mathbf{R}^2 . The drawing shows a 2-dimensional linear subspace of \mathbf{R}^3 and the corresponding line in the plane $x_3 = 1$. (The same works for affine subspaces of \mathbf{R}^d and linear subspaces of \mathbf{R}^{d+1} not contained in the subspace $x_{d+1} = 0$.)

This correspondence also leads directly to extending the affine plane \mathbf{R}^2 into the *projective plane*: To the points of \mathbf{R}^2 corresponding to nonhorizontal

lines through 0 in \mathbf{R}^3 we add points “at infinity,” that correspond to horizontal lines through 0 in \mathbf{R}^3 . But in this book we remain in the affine space most of the time, and we do not use the projective notions.

Let a_1, a_2, \dots, a_{d+1} be points in \mathbf{R}^d , and let A be the $d \times d$ matrix with $a_i - a_{d+1}$ as the i th column, $i = 1, 2, \dots, d$. Then a_1, \dots, a_{d+1} are affinely independent if and only if A has d linearly independent columns, and this is equivalent to $\det(A) \neq 0$. We have a useful criterion of affine independence using a determinant.

Affine subspaces of \mathbf{R}^d of certain dimensions have special names. A $(d-1)$ -dimensional affine subspace of \mathbf{R}^d is called a *hyperplane* (while the word *plane* usually means a 2-dimensional subspace of \mathbf{R}^d for any d). One-dimensional subspaces are lines, and a k -dimensional affine subspace is often called a *k-flat*.

A hyperplane is usually specified by a single linear equation of the form $a_1x_1 + a_2x_2 + \dots + a_dx_d = b$. We usually write the left-hand side as the scalar product $\langle a, x \rangle$. So a hyperplane can be expressed as the set $\{x \in \mathbf{R}^d: \langle a, x \rangle = b\}$ where $a \in \mathbf{R}^d \setminus \{0\}$ and $b \in \mathbf{R}$. A (closed) *half-space* in \mathbf{R}^d is a set of the form $\{x \in \mathbf{R}^d: \langle a, x \rangle \geq b\}$ for some $a \in \mathbf{R}^d \setminus \{0\}$ and some $b \in \mathbf{R}$; the hyperplane $\{x \in \mathbf{R}^d: \langle a, x \rangle = b\}$ is its boundary.

General k -flats can be given either as intersections of hyperplanes or as affine images of \mathbf{R}^k (parametric expression). In the first case, an intersection of k hyperplanes can also be viewed as a solution to a system $Ax = b$ of linear equations, where $x \in \mathbf{R}^d$ is regarded as a column vector, A is a $k \times d$ matrix, and $b \in \mathbf{R}^k$. (As a rule, in formulas involving matrices, we interpret points of \mathbf{R}^d as *column* vectors.)

An *affine mapping* $f: \mathbf{R}^k \rightarrow \mathbf{R}^d$ has the form $f: y \mapsto By + c$ for some $d \times k$ matrix B and some $c \in \mathbf{R}^d$, so it is a composition of a linear map with a translation. The image of f is a k' -flat for some $k' \leq \min(k, d)$. This k' equals the rank of the matrix B .

General position. “We assume that the points (lines, hyperplanes, . . .) are in general position.” This magical phrase appears in many proofs. Intuitively, general position means that no “unlikely coincidences” happen in the considered configuration. For example, if 3 points are chosen in the plane without any special intention, “randomly,” they are unlikely to lie on a common line. For a planar point set in general position, we always require that no three of its points be collinear. For points in \mathbf{R}^d in general position, we assume similarly that no unnecessary affine dependencies exist: No $k \leq d+1$ points lie in a common $(k-2)$ -flat. For lines in the plane in general position, we postulate that no 3 lines have a common point and no 2 are parallel.

The precise meaning of general position is not fully standard: It may depend on the particular context, and to the usual conditions mentioned above we sometimes add others where convenient. For example, for a planar point set in general position we can also suppose that no two points have the same x -coordinate.

What conditions are suitable for including into a “general position” assumption? In other words, what can be considered as an unlikely coincidence? For example, let X be an n -point set in the plane, and let the coordinates of the i th point be (x_i, y_i) . Then the vector $v(X) = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n)$ can be regarded as a point of \mathbf{R}^{2n} . For a configuration X in which $x_1 = x_2$, i.e., the first and second points have the same x -coordinate, the point $v(X)$ lies on the hyperplane $\{x_1 = x_2\}$ in \mathbf{R}^{2n} . The configurations X where *some* two points share the x -coordinate thus correspond to the union of $\binom{n}{2}$ hyperplanes in \mathbf{R}^{2n} . Since a hyperplane in \mathbf{R}^{2n} has $(2n-1)$ -dimensional measure zero, almost all points of \mathbf{R}^{2n} correspond to planar configurations X with all the points having distinct x -coordinates. In particular, if X is any n -point planar configuration and $\varepsilon > 0$ is any given real number, then there is a configuration X' , obtained from X by moving each point by distance at most ε , such that all points of X' have distinct x -coordinates. Not only that: Almost all small movements (*perturbations*) of X result in X' with this property.

This is the key property of general position: Configurations in general position lie arbitrarily close to any given configuration (and they abound in any small neighborhood of any given configuration). Here is a fairly general type of condition with this property. Suppose that a configuration X is specified by a vector $t = (t_1, t_2, \dots, t_m)$ of m real numbers (coordinates). The objects of X can be points in \mathbf{R}^d , in which case $m = dn$ and the t_j are the coordinates of the points, but they can also be circles in the plane, with $m = 3n$ and the t_j expressing the center and the radius of each circle, and so on. The general position condition we can put on the configuration X is $p(t) = p(t_1, t_2, \dots, t_m) \neq 0$, where p is some nonzero polynomial in m variables. Here we use the following well-known fact (a consequence of Sard’s theorem; see, e.g., Bredon [Bre93], Appendix C): *For any nonzero m -variate polynomial $p(t_1, \dots, t_m)$, the zero set $\{t \in \mathbf{R}^m: p(t) = 0\}$ has measure 0 in \mathbf{R}^m .*

Therefore, almost all configurations X satisfy $p(t) \neq 0$. So any condition that can be expressed as $p(t) \neq 0$ for a certain polynomial p in m real variables, or, more generally, as $p_1(t) \neq 0$ or $p_2(t) \neq 0$ or \dots , for finitely or countably many polynomials p_1, p_2, \dots , can be included in a general position assumption.

For example, let X be an n -point set in \mathbf{R}^d , and let us consider the condition “no $d+1$ points of X lie in a common hyperplane.” In other words, no $d+1$ points should be affinely dependent. As we know, the affine dependence of $d+1$ points means that a suitable $d \times d$ determinant equals 0. This determinant is a polynomial (of degree d) in the coordinates of these $d+1$ points. Introducing one polynomial for every $(d+1)$ -tuple of the points, we obtain $\binom{n}{d+1}$ polynomials such that at least one of them is 0 for any configuration X with $d+1$ points in a common hyperplane. Other usual conditions for general position can be expressed similarly.

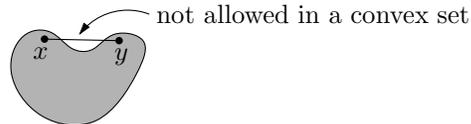
In many proofs, assuming general position simplifies matters considerably. But what do we do with configurations X_0 that are not in general position? We have to argue, somehow, that if the statement being proved is valid for configurations X arbitrarily close to our X_0 , then it must be valid for X_0 itself, too. Such proofs, usually called *perturbation arguments*, are often rather simple, and almost always somewhat boring. But sometimes they can be tricky, and one should not underestimate them, no matter how tempting this may be. A nontrivial example will be demonstrated in Section 5.5 (Lemma 5.5.4).

Exercises

1. Verify that the affine hull of a set $X \subseteq \mathbf{R}^d$ equals the set of all affine combinations of points of X . [2]
2. Let A be a 2×3 matrix and let $b \in \mathbf{R}^2$. Interpret the solution of the system $Ax = b$ geometrically (in most cases, as an intersection of two planes) and discuss the possible cases in algebraic and geometric terms. [2]
3. (a) What are the possible intersections of two (2-dimensional) planes in \mathbf{R}^4 ? What is the “typical” case (general position)? What about two hyperplanes in \mathbf{R}^4 ? [3]
 (b) Objects in \mathbf{R}^4 can sometimes be “visualized” as objects in \mathbf{R}^3 moving in time (so time is interpreted as the fourth coordinate). Try to visualize the intersection of two planes in \mathbf{R}^4 discussed (a) in this way.

1.2 Convex Sets, Convex Combinations, Separation

Intuitively, a set is convex if its surface has no “dips”:



1.2.1 Definition (Convex set). A set $C \subseteq \mathbf{R}^d$ is convex if for every two points $x, y \in C$ the whole segment xy is also contained in C . In other words, for every $t \in [0, 1]$, the point $tx + (1 - t)y$ belongs to C .

The intersection of an arbitrary family of convex sets is obviously convex. So we can define the *convex hull* of a set $X \subseteq \mathbf{R}^d$, denoted by $\text{conv}(X)$, as the intersection of all convex sets in \mathbf{R}^d containing X . Here is a planar example with a finite X :



An alternative description of the convex hull can be given using convex combinations.

1.2.2 Claim. *A point x belongs to $\text{conv}(X)$ if and only if there exist points $x_1, x_2, \dots, x_n \in X$ and nonnegative real numbers t_1, t_2, \dots, t_n with $\sum_{i=1}^n t_i = 1$ such that $x = \sum_{i=1}^n t_i x_i$.*

The expression $\sum_{i=1}^n t_i x_i$ as in the claim is called a *convex combination* of the points x_1, x_2, \dots, x_n . (Compare this with the definitions of linear and affine combinations.)

Sketch of proof. Each convex combination of points of X must lie in $\text{conv}(X)$: For $n = 2$ this is by definition, and for larger n by induction. Conversely, the set of all convex combinations obviously contains X , and it is convex. \square

In \mathbf{R}^d , it is sufficient to consider convex combinations involving at most $d+1$ points:

1.2.3 Theorem (Carathéodory's theorem). *Let $X \subseteq \mathbf{R}^d$. Then each point of $\text{conv}(X)$ is a convex combination of at most $d+1$ points of X .*

For example, in the plane, $\text{conv}(X)$ is the union of all triangles with vertices at points of X . The proof of the theorem is left as an exercise to the subsequent section.

A basic result about convex sets is the separability of disjoint convex sets by a hyperplane.

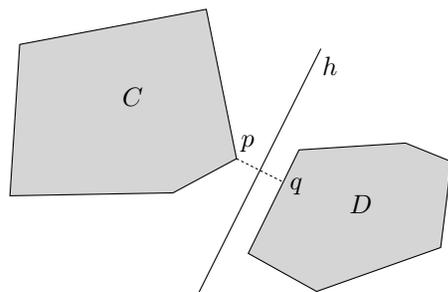
1.2.4 Theorem (Separation theorem). *Let $C, D \subseteq \mathbf{R}^d$ be convex sets with $C \cap D = \emptyset$. Then there exists a hyperplane h such that C lies in one of the closed half-spaces determined by h , and D lies in the opposite closed half-space. In other words, there exist a unit vector $a \in \mathbf{R}^d$ and a number $b \in \mathbf{R}$ such that for all $x \in C$ we have $\langle a, x \rangle \geq b$, and for all $x \in D$ we have $\langle a, x \rangle \leq b$.*

If C and D are closed and at least one of them is bounded, they can be separated strictly; in such a way that $C \cap h = D \cap h = \emptyset$.

In particular, a closed convex set can be strictly separated from a point. This implies that the convex hull of a closed set X equals the intersection of all closed half-spaces containing X .

Sketch of proof. First assume that C and D are compact (i.e., closed and bounded). Then the Cartesian product $C \times D$ is a compact space, too, and the distance function $(x, y) \mapsto \|x - y\|$ attains its minimum on $C \times D$. That is, there exist points $p \in C$ and $q \in D$ such that the distance of C and D equals the distance of p and q .

The desired separating hyperplane h can be taken as the one perpendicular to the segment pq and passing through its midpoint:



It is easy to check that h indeed avoids both C and D .

If D is compact and C closed, we can intersect C with a large ball and get a compact set C' . If the ball is sufficiently large, then C and C' have the same distance to D . So the distance of C and D is attained at some $p \in C'$ and $q \in D$, and we can use the previous argument.

For arbitrary disjoint convex sets C and D , we choose a sequence $C_1 \subseteq C_2 \subseteq C_3 \subseteq \dots$ of compact convex subsets of C with $\bigcup_{n=1}^{\infty} C_n = C$. For example, assuming that $0 \in C$, we can let C_n be the intersection of the closure of $(1 - \frac{1}{n})C$ with the ball of radius n centered at 0. A similar sequence $D_1 \subseteq D_2 \subseteq \dots$ is chosen for D , and we let $h_n = \{x \in \mathbf{R}^d: \langle a_n, x \rangle = b_n\}$ be a hyperplane separating C_n from D_n , where a_n is a unit vector and $b_n \in \mathbf{R}$. The sequence $(b_n)_{n=1}^{\infty}$ is bounded, and by compactness, the sequence of $(d+1)$ -component vectors $(a_n, b_n) \in \mathbf{R}^{d+1}$ has a cluster point (a, b) . One can verify, by contradiction, that the hyperplane $h = \{x \in \mathbf{R}^d: \langle a, x \rangle = b\}$ separates C and D (nonstrictly). \square

The importance of the separation theorem is documented by its presence in several branches of mathematics in various disguises. Its home territory is probably functional analysis, where it is formulated and proved for infinite-dimensional spaces; essentially it is the so-called Hahn–Banach theorem. The usual functional-analytic proof is different from the one we gave, and in a way it is more elegant and conceptual. The proof sketched above uses more special properties of \mathbf{R}^d , but it is quite short and intuitive in the case of compact C and D .

Connection to linear programming. A basic result in the theory of linear programming is the Farkas lemma. It is a special case of the duality of linear programming (discussed in Section ??) as well as the key step in its proof.

1.2.5 Lemma (Farkas lemma, one of many versions). *For every $d \times n$ real matrix A , exactly one of the following cases occurs:*

- (i) *The system of linear equations $Ax = 0$ has a nontrivial nonnegative solution $x \in \mathbf{R}^n$ (all components of x are nonnegative and at least one of them is strictly positive).*

- (ii) There exists a $y \in \mathbf{R}^d$ such that $y^T A$ is a vector with all entries strictly negative. Thus, if we multiply the j th equation in the system $Ax = 0$ by y_j and add these equations together, we obtain an equation that obviously has no nontrivial nonnegative solution, since all the coefficients on the left-hand sides are strictly negative, while the right-hand side is 0.

Proof. Let us see why this is yet another version of the separation theorem. Let $V \subset \mathbf{R}^d$ be the set of n points given by the column vectors of the matrix A . We distinguish two cases: Either $0 \in \text{conv}(V)$ or $0 \notin \text{conv}(V)$.

In the former case, we know that 0 is a convex combination of the points of V , and the coefficients of this convex combination determine a nontrivial nonnegative solution to $Ax = 0$.

In the latter case, there exists a hyperplane strictly separating V from 0 , i.e., a unit vector $y \in \mathbf{R}^d$ such that $\langle y, v \rangle < \langle y, 0 \rangle = 0$ for each $v \in V$. This is just the y from the second alternative in the Farkas lemma. \square

Bibliography and remarks. Most of the material in this chapter is quite old and can be found in many surveys and textbooks. Providing historical accounts of such well-covered areas is not among the goals of this book, and so we mention only a few references for the specific results discussed in the text and add some remarks concerning related results.

The concept of convexity and the rudiments of convex geometry have been around since antiquity. The initial chapter of the *Handbook of Convex Geometry* [GW93] succinctly describes the history, and the handbook can be recommended as the basic source on questions related to convexity, although knowledge has progressed significantly since its publication.

For an introduction to functional analysis, including the Hahn–Banach theorem, see Rudin [Rud91], for example. The Farkas lemma originated in [Far94] (nineteenth century!). More on the history of the duality of linear programming can be found, e.g., in Schrijver’s book [Sch86].

As for the origins, generalizations, and applications of Carathéodory’s theorem, as well as of Radon’s lemma and Helly’s theorem discussed in the subsequent sections, a recommendable survey is Eckhoff [Eck93], and an older well-known source is Danzer, Grünbaum, and Klee [DGK63].

Carathéodory’s theorem comes from the paper [Car07], concerning power series and harmonic analysis. A somewhat similar theorem, due to Steinitz [Ste16], asserts that if x lies in the interior of $\text{conv}(X)$ for an $X \subseteq \mathbf{R}^d$, then it also lies in the interior of $\text{conv}(Y)$ for some $Y \subseteq X$ with $|Y| \leq 2d$. Bonnice and Klee [BK63] proved a common generalization of both these theorems: Any k -interior point of $\text{conv}(X)$ is a k -interior point of $\text{conv}(Y)$ for some $Y \subseteq X$ with at most $\max(2k, d+1)$

points, where x is called a k -interior point of a set C if it lies in the relative interior of the convex hull of some $k+1$ affinely independent points of C .

Exercises

1. Give a detailed proof of Claim 1.2.2. \square
2. Write down a detailed proof of the separation theorem. \square
3. Find an example of two disjoint closed convex sets in the plane that are not strictly separable. \square
4. Let $f: \mathbf{R}^d \rightarrow \mathbf{R}^k$ be an affine map.
 - (a) Prove that if $C \subseteq \mathbf{R}^d$ is convex, then $f(C)$ is convex as well. Is the preimage of a convex set always convex? \square
 - (b) For $X \subseteq \mathbf{R}^d$ arbitrary, prove that $f(\text{conv}(X)) = \text{conv}(f(X))$. \square
5. Let $X \subseteq \mathbf{R}^d$. Prove that $\text{diam}(\text{conv}(X)) = \text{diam}(X)$, where the diameter $\text{diam}(Y)$ of a set Y is $\sup\{\|x - y\|: x, y \in Y\}$. \square
6. A set $C \subseteq \mathbf{R}^d$ is a *convex cone* if it is convex and for each $x \in C$, the ray $\vec{0x}$ is fully contained in C .
 - (a) Analogously to the convex and affine hulls, define the appropriate “conic hull” and the corresponding notion of “combination” (analogous to the convex and affine combinations). \square
 - (b) Let C be a convex cone in \mathbf{R}^d and $b \notin C$ a point. Prove that there exists a vector a with $\langle a, x \rangle \geq 0$ for all $x \in C$ and $\langle a, b \rangle < 0$. \square
7. (Variations on the Farkas lemma) Let A be a $d \times n$ matrix and let $b \in \mathbf{R}^d$.
 - (a) Prove that the system $Ax = b$ has a nonnegative solution $x \in \mathbf{R}^n$ if and only if every $y \in \mathbf{R}^d$ satisfying $y^T A \geq 0$ also satisfies $y^T b \geq 0$. \square
 - (b) Prove that the system of inequalities $Ax \leq b$ has a nonnegative solution x if and only if every nonnegative $y \in \mathbf{R}^d$ with $y^T A \geq 0$ also satisfies $y^T b \geq 0$. \square
8. (a) Let $C \subset \mathbf{R}^d$ be a compact convex set with a nonempty interior, and let $p \in C$ be an interior point. Show that there exists a line ℓ passing through p such that the segment $\ell \cap C$ is at least as long as any segment parallel to ℓ and contained in C . \square
 - (b) Show that (a) may fail for C compact but not convex. \square

1.3 Radon's Lemma and Helly's Theorem

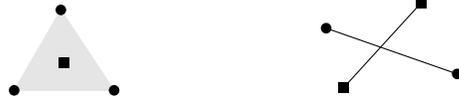
Carathéodory's theorem from the previous section, together with Radon's lemma and Helly's theorem presented here, are three basic properties of convexity in \mathbf{R}^d involving the dimension. We begin with Radon's lemma.

1.3.1 Theorem (Radon's lemma). *Let A be a set of $d+2$ points in \mathbf{R}^d . Then there exist two disjoint subsets $A_1, A_2 \subset A$ such that*

$$\text{conv}(A_1) \cap \text{conv}(A_2) = \emptyset.$$

A point $x \in \text{conv}(A_1) \cap \text{conv}(A_2)$, where A_1 and A_2 are as in the theorem, is called a *Radon point* of A , and the pair (A_1, A_2) is called a *Radon partition* of A (it is easily seen that we can require $A_1 \cup A_2 = A$).

Here are two possible cases in the plane:



Proof. Let $A = \{a_1, a_2, \dots, a_{d+2}\}$. These $d+2$ points are necessarily affinely dependent. That is, there exist real numbers $\alpha_1, \dots, \alpha_{d+2}$, not all of them 0, such that $\sum_{i=1}^{d+2} \alpha_i = 0$ and $\sum_{i=1}^{d+2} \alpha_i a_i = 0$.

Set $P = \{i: \alpha_i > 0\}$ and $N = \{i: \alpha_i < 0\}$. Both P and N are nonempty. We claim that P and N determine the desired subsets. Let us put $A_1 = \{a_i: i \in P\}$ and $A_2 = \{a_i: i \in N\}$. We are going to exhibit a point x that is contained in the convex hulls of both these sets.

Put $S = \sum_{i \in P} \alpha_i$; we also have $S = -\sum_{i \in N} \alpha_i$. Then we define

$$x = \sum_{i \in P} \frac{\alpha_i}{S} a_i. \quad (1.1)$$

Since $\sum_{i=1}^{d+2} \alpha_i a_i = 0 = \sum_{i \in P} \alpha_i a_i + \sum_{i \in N} \alpha_i a_i$, we also have

$$x = \sum_{i \in N} \frac{-\alpha_i}{S} a_i. \quad (1.2)$$

The coefficients of the a_i in (1.1) are nonnegative and sum to 1, so x is a convex combination of points of A_1 . Similarly, (1.2) expresses x as a convex combination of points of A_2 . \square

Helly's theorem is one of the most famous results of a combinatorial nature about convex sets.

1.3.2 Theorem (Helly's theorem). *Let C_1, C_2, \dots, C_n be convex sets in \mathbf{R}^d , $n \geq d+1$. Suppose that the intersection of every $d+1$ of these sets is nonempty. Then the intersection of all the C_i is nonempty.*

The first nontrivial case states that if every 3 among 4 convex sets in the plane intersect, then there is a point common to all 4 sets. This can be proved by an elementary geometric argument, perhaps distinguishing a few cases, and the reader may want to try to find a proof before reading further.

In a contrapositive form, Helly's theorem guarantees that whenever C_1, C_2, \dots, C_n are convex sets with $\bigcap_{i=1}^n C_i = \emptyset$, then this is witnessed by some at most $d+1$ sets with empty intersection among the C_i . In this way, many proofs are greatly simplified, since in planar problems, say, one can deal with 3 convex sets instead of an arbitrary number, as is amply illustrated in the exercises below.

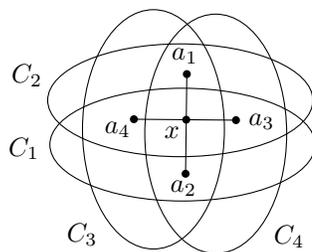
It is very tempting and quite usual to formulate Helly's theorem as follows: "If every $d+1$ among n convex sets in \mathbf{R}^d intersect, then all the sets intersect." But, strictly speaking, this is false, for a trivial reason: For $d \geq 2$, the assumption as stated here is met by $n = 2$ disjoint convex sets.

Proof of Helly's theorem. (Using Radon's lemma.) For a fixed d , we proceed by induction on n . The case $n = d+1$ is clear, so we suppose that $n \geq d+2$ and that the statement of Helly's theorem holds for smaller n . Actually, $n = d+2$ is the crucial case; the result for larger n follows at once by a simple induction.

Consider sets C_1, C_2, \dots, C_n satisfying the assumptions. If we leave out any one of these sets, the remaining sets have a nonempty intersection by the inductive assumption. Let us fix a point $a_i \in \bigcap_{j \neq i} C_j$ and consider the points a_1, a_2, \dots, a_{d+2} . By Radon's lemma, there exist disjoint index sets $I_1, I_2 \subset \{1, 2, \dots, d+2\}$ such that

$$\text{conv}(\{a_i: i \in I_1\}) \cap \text{conv}(\{a_i: i \in I_2\}) \neq \emptyset.$$

We pick a point x in this intersection. The following picture illustrates the case $d = 2$ and $n = 4$:



We claim that x lies in the intersection of all the C_i . Consider some $i \in \{1, 2, \dots, n\}$; then $i \notin I_1$ or $i \notin I_2$. In the former case, each a_j with $j \in I_1$ lies in C_i , and so $x \in \text{conv}(\{a_j: j \in I_1\}) \subseteq C_i$. For $i \notin I_2$ we similarly conclude that $x \in \text{conv}(\{a_j: j \in I_2\}) \subseteq C_i$. Therefore, $x \in \bigcap_{i=1}^n C_i$. \square

An infinite version of Helly's theorem. If we have an infinite collection of convex sets in \mathbf{R}^d such that any $d+1$ of them have a common point, the entire collection still need not have a common point. Two examples in \mathbf{R}^1 are the families of intervals $\{(0, 1/n): n = 1, 2, \dots\}$ and $\{[n, \infty): n = 1, 2, \dots\}$. The sets in the first example are not closed, and the second example uses unbounded sets. For *compact* (i.e., closed and bounded) sets, the theorem holds:

1.3.3 Theorem (Infinite version of Helly's theorem). *Let \mathcal{C} be an arbitrary infinite family of compact convex sets in \mathbf{R}^d such that any $d+1$ of the sets have a nonempty intersection. Then all the sets of \mathcal{C} have a nonempty intersection.*

Proof. By Helly's theorem, any finite subfamily of \mathcal{C} has a nonempty intersection. By a basic property of compactness, if we have an arbitrary family of compact sets such that each of its finite subfamilies has a nonempty intersection, then the entire family has a nonempty intersection. \square

Several nice applications of Helly's theorem are indicated in the exercises below, and we will meet a few more later in this book.

Bibliography and remarks. Helly proved Theorem 1.3.2 in 1913 and communicated it to Radon, who published a proof in [Rad21]. This proof uses Radon's lemma, although the statement wasn't explicitly formulated in Radon's paper. References to many other proofs and generalizations can be found in the already mentioned surveys [Eck93] and [DGK63].

Helly's theorem inspired a whole industry of Helly-type theorems. A family \mathcal{B} of sets is said to have *Helly number* h if the following holds: Whenever a finite subfamily $\mathcal{F} \subseteq \mathcal{B}$ is such that every h or fewer sets of \mathcal{F} have a common point, then $\bigcap \mathcal{F} \neq \emptyset$. So Helly's theorem says that the family of all convex sets in \mathbf{R}^d has Helly number $d+1$. More generally, let P be some property of families of sets that is hereditary, meaning that if \mathcal{F} has property P and $\mathcal{F}' \subseteq \mathcal{F}$, then \mathcal{F}' has P as well. A family \mathcal{B} is said to have Helly number h with respect to P if for every finite $\mathcal{F} \subseteq \mathcal{B}$, all subfamilies of \mathcal{F} of size at most h having P implies \mathcal{F} having P . That is, the absence of P is always witnessed by some at most h sets, so it is a "local" property.

Exercises

1. Prove Carathéodory's theorem (you may use Radon's lemma). \square
2. Let $K \subset \mathbf{R}^d$ be a convex set and let $C_1, \dots, C_n \subseteq \mathbf{R}^d$, $n \geq d+1$, be convex sets such that the intersection of every $d+1$ of them contains a translated copy of K . Prove that then the intersection of all the sets C_i also contains a translated copy of K . \square
This result was noted by Vincensini [Vin39] and by Klee [Kle53].
3. Find an example of 4 convex sets in the plane such that the intersection of each 3 of them contains a segment of length 1, but the intersection of all 4 contains no segment of length 1. \square
4. A *strip of width* w is a part of the plane bounded by two parallel lines at distance w . The *width* of a set $X \subseteq \mathbf{R}^2$ is the smallest width of a strip containing X .
 - (a) Prove that a compact convex set of width 1 contains a segment of length 1 of every direction. \square
 - (b) Let $\{C_1, C_2, \dots, C_n\}$ be closed convex sets in the plane, $n \geq 3$, such that the intersection of every 3 of them has width at least 1. Prove that $\bigcap_{i=1}^n C_i$ has width at least 1. \square

The result as in (b), for arbitrary dimension d , was proved by Sallee [Sal75], and a simple argument using Helly's theorem was noted by Buchman and Valentine [BV82].

5. Statement: Each set $X \subset \mathbf{R}^2$ of diameter at most 1 (i.e., any 2 points have distance at most 1) is contained in some disc of radius $1/\sqrt{3}$.
- (a) Prove the statement for 3-element sets X . [2]
 - (b) Prove the statement for all finite sets X . [2]
 - (c) Generalize the statement to \mathbf{R}^d : determine the smallest $r = r(d)$ such that every set of diameter 1 in \mathbf{R}^d is contained in a ball of radius r (prove your claim). [4]

The result as in (c) is due to Jung; see [DGK63].

6. Let $C \subset \mathbf{R}^d$ be a compact convex set. Prove that the mirror image of C can be covered by a suitable translate of C blown up by the factor of d ; that is, there is an $x \in \mathbf{R}^d$ with $-C \subseteq x + dC$. [4]
7. (a) Prove that if the intersection of each 4 or fewer among convex sets $C_1, \dots, C_n \subseteq \mathbf{R}^2$ contains a ray then $\bigcap_{i=1}^n C_i$ also contains a ray. [4]
- (b) Show that the number 4 in (a) cannot be replaced by 3. [2]
- This result, and an analogous one in \mathbf{R}^d with the Helly number $2d$, are due to Katchalski [Kat78].

8. For a set $X \subseteq \mathbf{R}^2$ and a point $x \in X$, let us denote by $V(x)$ the set of all points $y \in X$ that can "see" x , i.e., points such that the segment xy is contained in X . The *kernel* of X is defined as the set of all points $x \in X$ such that $V(x) = X$. A set with a nonempty kernel is called *star-shaped*.
- (a) Prove that the kernel of any set is convex. [1]
 - (b) Prove that if $V(x) \cap V(y) \cap V(z) \neq \emptyset$ for every $x, y, z \in X$ and X is compact, then X is star-shaped. That is, if every 3 paintings in a (planar) art gallery can be seen at the same time from some location (possibly different for different triples of paintings), then all paintings can be seen simultaneously from somewhere. If it helps, assume that X is a polygon. [5]

- (c) Construct a nonempty set $X \subseteq \mathbf{R}^2$ such that each of its finite subsets can be seen from some point of X but X is not star-shaped. [2]

The result in (b), as well as the d -dimensional generalization (with every $d+1$ regions $V(x)$ intersecting), is called Krasnosel'skii's theorem; see [Eck93] for references and related results.

9. In the situation of Radon's lemma (A is a $(d+2)$ -point set in \mathbf{R}^d), call a point $x \in \mathbf{R}^d$ a *Radon point* of A if it is contained in convex hulls of two disjoint subsets of A . Prove that if A is in general position (no $d+1$ points affinely dependent), then its Radon point is unique. [3]
10. (a) Let $X, Y \subset \mathbf{R}^2$ be finite point sets, and suppose that for every subset $S \subseteq X \cup Y$ of at most 4 points, $S \cap X$ can be separated (strictly) by a line from $S \cap Y$. Prove that X and Y are line-separable. [3]
- (b) Extend (a) to sets $X, Y \subset \mathbf{R}^d$, with $|S| \leq d+2$. [5]
- The result (b) is called Kirchberger's theorem [Kir03].

1.4 Centerpoint and Ham Sandwich

We prove an interesting result as an application of Helly's theorem.

1.4.1 Definition (Centerpoint). Let X be an n -point set in \mathbf{R}^d . A point $x \in \mathbf{R}^d$ is called a centerpoint of X if each closed half-space containing x contains at least $\frac{n}{d+1}$ points of X .

Let us stress that one set may generally have many centerpoints, and a centerpoint need not belong to X .

The notion of centerpoint can be viewed as a generalization of the *median* of one-dimensional data. Suppose that $x_1, \dots, x_n \in \mathbf{R}$ are results of measurements of an unknown real parameter x . How do we estimate x from the x_i ? We can use the arithmetic mean, but if one of the measurements is completely wrong (say, 100 times larger than the others), we may get quite a bad estimate. A more "robust" estimate is a *median*, i.e., a point x such that at least $\frac{n}{2}$ of the x_i lie in the interval $(-\infty, x]$ and at least $\frac{n}{2}$ of them lie in $[x, \infty)$. The centerpoint can be regarded as a generalization of the median for higher-dimensional data.

In the definition of centerpoint we could replace the fraction $\frac{1}{d+1}$ by some other parameter $\alpha \in (0, 1)$. For $\alpha > \frac{1}{d+1}$, such an " α -centerpoint" need not always exist: Take $d+1$ points in general position for X . With $\alpha = \frac{1}{d+1}$ as in the definition above, a centerpoint always exists, as we prove next.

Centerpoints are important, for example, in some algorithms of divide-and-conquer type, where they help divide the considered problem into smaller subproblems. Since no really efficient algorithms are known for finding "exact" centerpoints, the algorithms often use α -centerpoints with a suitable $\alpha < \frac{1}{d+1}$, which are easier to find.

1.4.2 Theorem (Centerpoint theorem). Each finite point set in \mathbf{R}^d has at least one centerpoint.

Proof. First we note an *equivalent definition of a centerpoint*: x is a centerpoint of X if and only if it lies in each open half-space γ such that $|X \cap \gamma| > \frac{d}{d+1}n$.

We would like to apply Helly's theorem to conclude that all these open half-spaces intersect. But we cannot proceed directly, since we have infinitely many half-spaces and they are open and unbounded. Instead of such an open half-space γ , we thus consider the compact convex set $\text{conv}(X \cap \gamma) \subset \gamma$.



Letting γ run through all open half-spaces γ with $|X \cap \gamma| > \frac{d}{d+1}n$, we obtain a family \mathcal{C} of compact convex sets. Each of them contains more than $\frac{d}{d+1}n$ points of X , and so the intersection of any $d+1$ of them contains at least one point of X . The family \mathcal{C} consists of finitely many distinct sets (since X has finitely many distinct subsets), and so $\bigcap \mathcal{C} \neq \emptyset$ by Helly's theorem. Each point in this intersection is a centerpoint. \square

In the definition of a centerpoint we can regard the finite set X as defining a distribution of mass in \mathbf{R}^d . The centerpoint theorem asserts that for some point x , any half-space containing x encloses at least $\frac{1}{d+1}$ of the total mass. It is not difficult to show that this remains valid for continuous mass distributions, or even for arbitrary Borel probability measures on \mathbf{R}^d (Exercise 1).

Ham-sandwich theorem and its relatives. Here is another important result, not much related to convexity but with a flavor resembling the centerpoint theorem.

1.4.3 Theorem (Ham-sandwich theorem). *Every d finite sets in \mathbf{R}^d can be simultaneously bisected by a hyperplane. A hyperplane h bisects a finite set A if each of the open half-spaces defined by h contains at most $\lfloor |A|/2 \rfloor$ points of A .*

This theorem is usually proved via continuous mass distributions using a tool from algebraic topology: the *Borsuk–Ulam theorem*. Here we omit a proof.

Note that if A_i has an odd number of points, then every h bisecting A_i passes through a point of A_i . Thus if A_1, \dots, A_d all have odd sizes and their union is in general position, then every hyperplane simultaneously bisecting them is determined by d points, one of each A_i . In particular, there are only finitely many such hyperplanes.

Again, an analogous ham-sandwich theorem holds for arbitrary d Borel probability measures in \mathbf{R}^d .

Center transversal theorem. There can be beautiful new things to discover even in well-studied areas of mathematics. A good example is the following recent result, which “interpolates” between the centerpoint theorem and the ham-sandwich theorem.

1.4.4 Theorem (Center transversal theorem). *Let $1 \leq k \leq d$ and let A_1, A_2, \dots, A_k be finite point sets in \mathbf{R}^d . Then there exists a $(k-1)$ -flat f such that for every hyperplane h containing f , both the closed half-spaces defined by h contain at least $\frac{1}{d-k+2}|A_i|$ points of A_i , $i = 1, 2, \dots, k$.*

The ham-sandwich theorem is obtained for $k = d$ and the centerpoint theorem for $k = 1$. The proof, which we again have to omit, is based on a result of algebraic topology, too, but it uses a considerably more advanced

machinery than the ham-sandwich theorem. However, the weaker result with $\frac{1}{d+1}$ instead of $\frac{1}{d-k+2}$ is easy to prove; see Exercise 2.

Bibliography and remarks. The centerpoint theorem was established by Rado [Rad47]. According to Steinlein's survey [Ste85], the ham-sandwich theorem was conjectured by Steinhaus (who also invented the popular 3-dimensional interpretation, namely, that the ham, the cheese, and the bread in any ham sandwich can be simultaneously bisected by a single straight motion of the knife) and proved by Banach. The center transversal theorem was found by Doľnikov [Doľ92] and, independently, by Živaljević and Vrećica [ŽV90].

Significant effort has been devoted to efficient algorithms for finding (approximate) centerpoints and ham-sandwich cuts (i.e., hyperplanes as in the ham-sandwich theorem). In the plane, a ham-sandwich cut for two n -point sets can be computed in linear time (Lo, Matoušek, and Steiger [LMS94]). In a higher but fixed dimension, the complexity of the best exact algorithms is currently slightly better than $O(n^{d-1})$. A centerpoint in the plane, too, can be found in linear time (Jadhav and Mukhopadhyay [JM94]). Both approximate ham-sandwich cuts (in the ratio $1 : 1+\varepsilon$ for a fixed $\varepsilon > 0$) and approximate centerpoints ($(\frac{1}{d+1}-\varepsilon)$ -centerpoints) can be computed in time $O(n)$ for every fixed dimension d and every fixed $\varepsilon > 0$, but the constant depends exponentially on d , and the algorithms are impractical if the dimension is not quite small. A practically efficient randomized algorithm for computing approximate centerpoints in high dimensions (α -centerpoints with $\alpha \approx 1/d^2$) was given by Clarkson, Eppstein, Miller, Sturtivant, and Teng [CEM⁺96].

Exercises

- (Centerpoints for general mass distributions)
 - Let μ be a Borel probability measure on \mathbf{R}^d ; that is, $\mu(\mathbf{R}^d) = 1$ and each open set is measurable. Show that for each open half-space γ with $\mu(\gamma) > t$ there exists a compact set $C \subset \gamma$ with $\mu(C) > t$. \square
 - Prove that each Borel probability measure in \mathbf{R}^d has a centerpoint (use (a) and the infinite Helly's theorem). \square
- Prove that for any k finite sets $A_1, \dots, A_k \subset \mathbf{R}^d$, where $1 \leq k \leq d$, there exists a $(k-1)$ -flat such that every hyperplane containing it has at least $\frac{1}{d+1} |A_i|$ points of A_i in both of its closed half-spaces for all $i = 1, 2, \dots, k$. \square

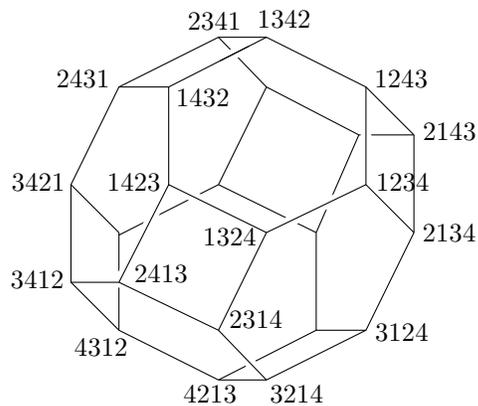
5

Convex Polytopes

Convex polytopes are convex hulls of finite point sets in \mathbf{R}^d . They constitute the most important class of convex sets with an enormous number of applications and connections.

Three-dimensional convex polytopes, especially the regular ones, have been fascinating people since the antiquity. Their investigation was one of the main sources of the theory of planar graphs, and thanks to this well-developed theory they are quite well understood. But convex polytopes in dimension 4 and higher are considerably more challenging, and a surprisingly deep theory, mainly of algebraic nature, was developed in attempts to understand their structure.

A strong motivation for the study of convex polytopes comes from practically significant areas such as combinatorial optimization, linear programming, and computational geometry. Let us look at a simple example illustrating how polytopes can be associated with combinatorial objects. The 3-dimensional polytope in the picture



is called the *permutahedron*. Although it is 3-dimensional, it is most naturally defined as a subset of \mathbf{R}^4 , namely, the convex hull of the 24 vectors obtained by permuting the coordinates of the vector $(1, 2, 3, 4)$ in all possible ways. In the picture, the (visible) vertices are labeled by the corresponding permutations. Similarly, the d -dimensional permutahedron is the convex hull of the $(d+1)!$ vectors in \mathbf{R}^{d+1} arising by permuting the coordinates of $(1, 2, \dots, d+1)$. One can observe that the edges of the polytope connect exactly pairs of permutations differing by a transposition of two adjacent numbers, and a closer examination reveals other connections between the structure of the permutahedron and properties of permutations.

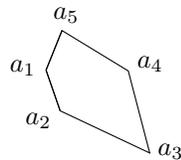
There are many other, more sophisticated, examples of convex polytopes assigned to combinatorial and geometric objects such as graphs, partially ordered sets, classes of metric spaces, or triangulations of a given point set. In many cases, such convex polytopes are a key tool for proving hard theorems about the original objects or for obtaining efficient algorithms. Two impressive examples are discussed in Chapter ??, and several others are scattered in other chapters.

The present chapter should convey some initial working knowledge of convex polytopes for a nonpolytopist. It is just a small sample of an extensive theory. A more comprehensive modern introduction is the book by Ziegler [Zie94].

5.1 Geometric Duality

First we discuss geometric duality, a simple technical tool indispensable in the study of convex polytopes and handy in many other situations. We begin with a simple motivating question.

How can we visualize the set of all lines intersecting a convex pentagon as in the picture?



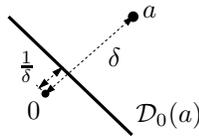
A suitable way is provided by line–point duality.

5.1.1 Definition (Duality transform). *The (geometric) duality transform is a mapping denoted by \mathcal{D}_0 . To a point $a \in \mathbf{R}^d \setminus \{0\}$ it assigns the hyperplane*

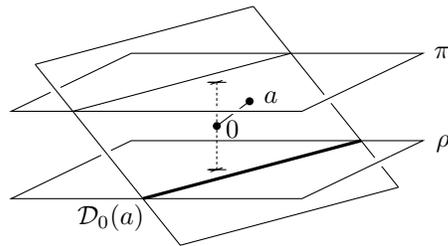
$$\mathcal{D}_0(a) = \{x \in \mathbf{R}^d: \langle a, x \rangle = 1\},$$

and to a hyperplane h not passing through the origin, which can be uniquely written in the form $h = \{x \in \mathbf{R}^d: \langle a, x \rangle = 1\}$, it assigns the point $\mathcal{D}_0(h) = a \in \mathbf{R}^d \setminus \{0\}$.

Here is the geometric meaning of the duality transform. If a is a point at distance δ from 0, then $\mathcal{D}_0(a)$ is the hyperplane perpendicular to the line $0a$ and intersecting that line at distance $\frac{1}{\delta}$ from 0, in the direction from 0 towards a .

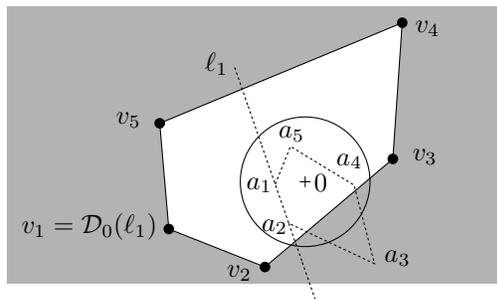


A nice interpretation of duality is obtained by working in \mathbf{R}^{d+1} and identifying the “primal” \mathbf{R}^d with the hyperplane $\pi = \{x \in \mathbf{R}^{d+1}: x_{d+1} = 1\}$ and the “dual” \mathbf{R}^d with the hyperplane $\rho = \{x \in \mathbf{R}^{d+1}: x_{d+1} = -1\}$. The hyperplane dual to a point $a \in \pi$ is produced as follows: We construct the hyperplane in \mathbf{R}^{d+1} perpendicular to $0a$ and containing 0, and we intersect it with ρ . Here is an illustration for $d = 2$:



In this way, the duality \mathcal{D}_0 can be naturally extended to k -flats in \mathbf{R}^d , whose duals are $(d-k-1)$ -flats. Namely, given a k -flat $f \subset \pi$, we consider the $(k+1)$ -flat F through 0 and f , we construct the orthogonal complement of F , and we intersect it with ρ , obtaining $\mathcal{D}_0(f)$.

Let us consider the pentagon drawn above and place it so that the origin lies in the interior. Let $v_i = \mathcal{D}_0(\ell_i)$, where ℓ_i is the line containing the side $a_i a_{i+1}$. Then the points dual to the lines intersecting the pentagon $a_1 a_2 \dots a_5$ fill exactly the exterior of the convex pentagon $v_1 v_2 \dots v_5$:



This follows easily from the properties of duality listed below (of course, there is nothing special about a pentagon here). Thus, the considered set of lines can be nicely described in the dual plane. A similar passage from lines to points or back is useful in many geometric or computational problems.

Properties of the duality transform. Let p be a point of \mathbf{R}^d distinct from the origin and let h be a hyperplane in \mathbf{R}^d not containing the origin. Let h^- stand for the closed half-space bounded by h and containing the origin, while h^+ denotes the other closed half-space bounded by h . That is, if $h = \{x \in \mathbf{R}^d: \langle a, x \rangle = 1\}$, then $h^- = \{x \in \mathbf{R}^d: \langle a, x \rangle \leq 1\}$.

5.1.2 Lemma (Duality preserves incidences).

- (i) $p \in h$ if and only if $\mathcal{D}_0(h) \in \mathcal{D}_0(p)$.
- (ii) $p \in h^-$ if and only if $\mathcal{D}_0(h) \in \mathcal{D}_0(p)^-$.

Proof. (i) Let $h = \{x \in \mathbf{R}^d: \langle a, x \rangle = 1\}$. Then $p \in h$ means $\langle a, p \rangle = 1$. Now, $\mathcal{D}_0(h)$ is the point a , and $\mathcal{D}_0(p)$ is the hyperplane $\{y \in \mathbf{R}^d: \langle y, p \rangle = 1\}$, and hence $\mathcal{D}_0(h) = a \in \mathcal{D}_0(p)$ also means just $\langle a, p \rangle = 1$. Part (ii) is proved similarly. \square

5.1.3 Definition (Dual set). For a set $X \subseteq \mathbf{R}^d$, we define the set dual to X , denoted by X^* , as follows:

$$X^* = \{y \in \mathbf{R}^d: \langle x, y \rangle \leq 1 \text{ for all } x \in X\}.$$

Another common name used for the duality is *polarity*; the dual set would then be called the *polar set*. Sometimes it is denoted by X° .

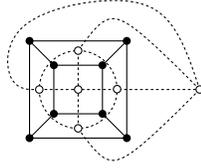
Geometrically, X^* is the intersection of all half-spaces of the form $\mathcal{D}_0(x)^-$ with $x \in X$. Or in other words, X^* consists of the origin plus all points y such that $X \subseteq \mathcal{D}_0(y)^-$. For example, if X is the pentagon $a_1 a_2 \dots a_5$ drawn above, then X^* is the pentagon $v_1 v_2 \dots v_5$.

For any set X , the set X^* is obviously closed and convex and contains the origin. Using the separation theorem (Theorem 1.2.4), it is easily shown that for any set $X \subseteq \mathbf{R}^d$, the set $(X^*)^*$ is the closure of $\text{conv}(X \cup \{0\})$. In particular, for a closed convex set containing the origin we have $(X^*)^* = X$ (Exercise 3).

For a hyperplane h , the dual set h^* is different from the point $\mathcal{D}_0(h)$.¹

For readers familiar with the duality of planar graphs, let us remark that it is closely related to the geometric duality applied to convex polytopes in \mathbf{R}^3 . For example, the next drawing illustrates a planar graph and its dual graph (dashed):

¹ In the literature, however, the “star” notation is sometimes also used for the dual point or hyperplane, so for a point p , the hyperplane $\mathcal{D}_0(p)$ would be denoted by p^* , and similarly, h^* may stand for $\mathcal{D}_0(h)$.



Later we will see that these are graphs of the 3-dimensional cube and of the regular octahedron, which are polytopes dual to each other in the sense defined above. A similar relation holds for all 3-dimensional polytopes and their graphs.

Other variants of duality. The duality transform \mathcal{D}_0 defined above is just one of a class of geometric transforms with similar properties. For some purposes, other such transforms (dualities) are more convenient. A particularly important duality, denoted by \mathcal{D} , corresponds to moving the origin to the “minus infinity” of the x_d -axis (the x_d -axis is considered vertical). A formal definition is as follows.

5.1.4 Definition (Another duality). *A nonvertical hyperplane h can be uniquely written in the form $h = \{x \in \mathbf{R}^d: x_d = a_1x_1 + \cdots + a_{d-1}x_{d-1} - a_d\}$. We set $\mathcal{D}(h) = (a_1, \dots, a_{d-1}, a_d)$. Conversely, the point $a = (a_1, \dots, a_{d-1}, a_d)$ maps back to h .*

The property (i) of Lemma 5.1.2 holds for this \mathcal{D} , and an analogue of (ii) is:

(ii') A point p lies above a hyperplane h if and only if the point $\mathcal{D}(h)$ lies above the hyperplane $\mathcal{D}(p)$.

Exercises

1. Let $C = \{x \in \mathbf{R}^d: |x_1| + \cdots + |x_d| \leq 1\}$. Show that C^* is the d -dimensional cube $\{x \in \mathbf{R}^d: \max |x_i| \leq 1\}$. Picture both bodies for $d = 3$. □
2. Prove the assertion made in the text about the lines intersecting a convex pentagon. □
3. Show that for any $X \subseteq \mathbf{R}^d$, $(X^*)^*$ equals the closure of $\text{conv}(X \cup \{0\})$, where X^* stands for the dual set to X . □
4. Let $C \subseteq \mathbf{R}^d$ be a convex set. Prove that C^* is bounded if and only if 0 lies in the interior of C . □
5. Show that $C = C^*$ if and only if C is the unit ball centered at the origin. □
6. (a) Let $C = \text{conv}(X) \subseteq \mathbf{R}^d$. Prove that $C^* = \bigcap_{x \in X} \mathcal{D}_0(x)^-$. □
 (b) Show that if $C = \bigcap_{h \in H} h^-$, where H is a collection of hyperplanes not passing through 0, and if C is bounded, then $C^* = \text{conv}\{\mathcal{D}_0(h): h \in H\}$. □
 (c) What is the right analogue of (b) if C is unbounded? □

7. What is the dual set h^* for a hyperplane h , and what about h^{**} ? \square
8. Verify the geometric interpretation of the duality \mathcal{D}_0 outlined in the text (using the embeddings of \mathbf{R}^d into \mathbf{R}^{d+1}). \square
9. (a) Let s be a segment in the plane. Describe the set of all points dual to lines intersecting s . \square
 (b) Consider $n \geq 3$ segments in the plane, such that none of them contains 0 but they all lie on lines passing through 0. Show that if every 3 among such segments can be intersected by a single line, then all the segments can be simultaneously intersected by a line. \square
 (c) Show that the assumption in (b) that the extensions of the segments pass through 0 is essential: For each $n \geq 1$, construct $n+1$ pairwise disjoint segments in the plane that cannot be simultaneously intersected by a line but every n of them can (such an example was first found by Hadwiger and Debrunner). \square

5.2 H -Polytopes and V -Polytopes

A convex polytope in the plane is a convex polygon. Famous examples of convex polytopes in \mathbf{R}^3 are the Platonic solids: regular tetrahedron, cube, regular octahedron, regular dodecahedron, and regular icosahedron. A convex polytope in \mathbf{R}^3 is a convex set bounded by finitely many convex polygons. Such a set can be regarded as a convex hull of a finite point set, or as an intersection of finitely many half-spaces. We thus define two types of convex polytopes, based on these two views.

5.2.1 Definition (H -polytope and V -polytope). *An H -polyhedron is an intersection of finitely many closed half-spaces in some \mathbf{R}^d . An H -polytope is a bounded H -polyhedron.*

A V -polytope is the convex hull of a finite point set in \mathbf{R}^d .

A basic theorem about convex polytopes claims that from the mathematical point of view, H -polytopes and V -polytopes are equivalent.

5.2.2 Theorem. *Each V -polytope is an H -polytope. Each H -polytope is a V -polytope.*

This is one of the theorems that may look “obvious” and whose proof needs no particularly clever idea but does require some work. In the present case, we do not intend to avoid it. Actually, we have quite a neat proof in store, but we postpone it to the end of this section.

Although H -polytopes and V -polytopes are mathematically equivalent, there is an enormous difference between them from the computational point of view. That is, it matters a lot whether a convex polytope is given to us as a convex hull of a finite set or as an intersection of half-spaces. For example, given a set of n points specifying a V -polytope, how do we find its representation as an H -polytope? It is not hard to come up with some

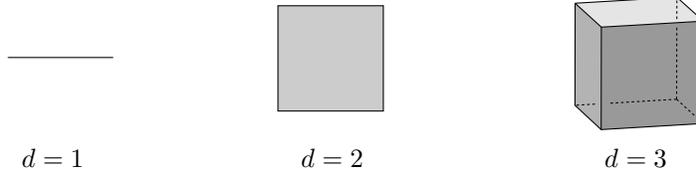
algorithm, but the problem is to find an efficient algorithm that would allow one to handle large real-world problems. This algorithmic question is not yet satisfactorily solved. Moreover, in some cases the number of required half-spaces may be astronomically large compared to the number n of points, as we will see later in this chapter.

As another illustration of the computational difference between *V*-polytopes and *H*-polytopes, we consider the maximization of a given linear function over a given polytope. For *V*-polytopes it is a trivial problem, since it suffices to substitute all points of V into the given linear function and select the maximum of the resulting values. But maximizing a linear function over the intersection of a collection of half-spaces is the basic problem of linear programming, and it is certainly nontrivial.

Terminology. The usual terminology does not distinguish *V*-polytopes and *H*-polytopes. A *convex polytope* means a point set in \mathbf{R}^d that is a *V*-polytope (and thus also an *H*-polytope). An arbitrary, possibly unbounded, *H*-polyhedron is called a *convex polyhedron*. All polytopes and polyhedra considered in this chapter are convex, and so the adjective “convex” is often omitted.

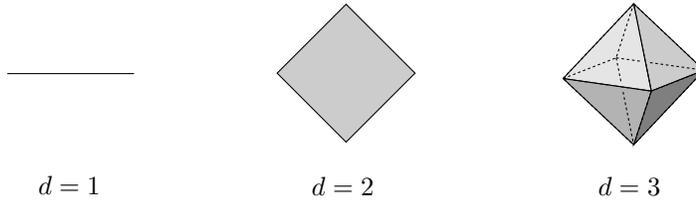
The *dimension* of a convex polyhedron is the dimension of its affine hull. It is the smallest dimension of a Euclidean space containing a congruent copy of P . The phrase “ d -dimensional polytope” is often abbreviated to “ d -polytope.”

Basic examples. One of the easiest classes of polytopes is that of *cubes*. The d -dimensional cube as a point set is the Cartesian product $[-1, 1]^d$.



As a *V*-polytope, the d -dimensional cube is the convex hull of the set $\{-1, 1\}^d$ (2^d points), and as an *H*-polytope, it can be described by the inequalities $-1 \leq x_i \leq 1$, $i = 1, 2, \dots, d$, i.e., by $2d$ half-spaces. We note that it is also the unit ball of the maximum norm $\|x\|_\infty = \max_i |x_i|$.

Another important example is the class of *crosspolytopes* (or generalized octahedra). The d -dimensional crosspolytope is the convex hull of the “coordinate cross,” i.e., $\text{conv}\{e_1, -e_1, e_2, -e_2, \dots, e_d, -e_d\}$, where e_1, \dots, e_d are the vectors of the standard orthonormal basis.



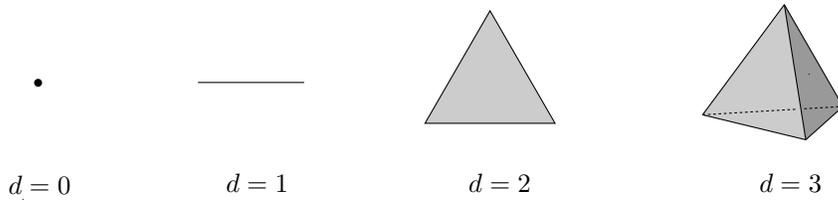
It is also the unit ball of the ℓ_1 -norm $\|x\|_1 = \sum_{i=1}^d |x_i|$. As an H -polytope, it can be expressed by the 2^d half-spaces of the form $\langle \sigma, x \rangle \leq 1$, where σ runs through all vectors in $\{-1, 1\}^d$.

The polytopes with the smallest possible number of vertices (for a given dimension) are called simplices.

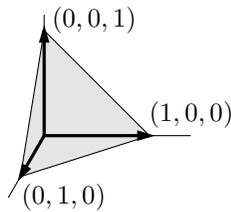
5.2.3 Definition (Simplex). A simplex is the convex hull of an affinely independent point set in some \mathbf{R}^d .

A d -dimensional simplex in \mathbf{R}^d can also be represented as an intersection of $d+1$ half-spaces, as is not difficult to check.

A *regular* d -dimensional simplex is the convex hull of $d+1$ points with all pairs of points having equal distances.



Unlike cubes and crosspolytopes, d -dimensional regular simplices do not have a very nice coordinate representation in \mathbf{R}^d . The simplest and most useful representation lives one dimension higher: The convex hull of the $d+1$ vectors e_1, \dots, e_{d+1} of the standard orthonormal basis in \mathbf{R}^{d+1} is a d -dimensional regular simplex with side length $\sqrt{2}$.

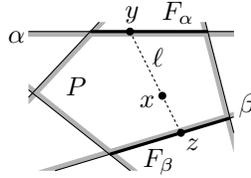


Proof of Theorem 5.2.2 (equivalence of H -polytopes and V -polytopes). We first show that any H -polytope is also a V -polytope. We proceed by induction on d . The case $d = 1$ being trivial, we suppose that $d \geq 2$.

So let Γ be a finite collection of closed half-spaces in \mathbf{R}^d such that $P = \bigcap \Gamma$ is nonempty and bounded. For each $\gamma \in \Gamma$, let $F_\gamma = P \cap \partial\gamma$ be the intersection of P with the bounding hyperplane of γ . Each nonempty F_γ is an H -polytope of dimension at most $d-1$ (correct?), and so it is the convex hull of a finite set $V_\gamma \subset F_\gamma$ by the inductive hypothesis.

We claim that $P = \text{conv}(V)$, where $V = \bigcup_{\gamma \in \Gamma} V_\gamma$. Let $x \in P$ and let ℓ be a line passing through x . The intersection $\ell \cap P$ is a segment; let y and z be its endpoints. There are $\alpha, \beta \in \Gamma$ such that $y \in F_\alpha$ and $z \in F_\beta$ (if y were

not on the boundary of any $\gamma \in \Gamma$, we could continue along ℓ a little further within P).



We have $y \in \text{conv}(V_\alpha)$ and $z \in \text{conv}(V_\beta)$, and thus $x \in \text{conv}(V_\alpha \cup V_\beta) \subseteq \text{conv}(V)$.

We have proved that any H -polytope is a V -polytope, and it remains to show that a V -polytope can be expressed as the intersection of finitely many half-spaces. This follows easily by duality (and implicitly uses the separation theorem).

Let $P = \text{conv}(V)$ with V finite, and assume that 0 is an interior point of P . By Exercise 5.1.6(a), the dual body P^* equals $\bigcap_{v \in V} \mathcal{D}_0(v)^-$, and by Exercise 5.1.4 it is bounded. By what we have already proved, P^* is a V -polytope, and by Exercise 5.1.6(a) again, $P = (P^*)^*$ is the intersection of finitely many half-spaces. \square

Bibliography and remarks. The theory of convex polytopes is a well-developed area covered in numerous books and surveys, such as the already recommended recent monograph [Zie94] (with addenda and updates on the web page of its author), the very influential book by Grünbaum whose new edition [Grü03] is accompanied by surveys of recent developments, the chapters on polytopes in the handbooks of discrete and computational geometry [GO97], of convex geometry [GW93], and of combinatorics [GGL95], or the books McMullen and Shephard [MS71] and Brønsted [Brø83], concentrating on questions about the numbers of faces. Recent progress in combinatorial and computational polytope theory is reflected in the collection [KZ00]. For analyzing examples, one should be aware of (free) software systems for manipulating convex polytopes, such as `polymake` by Gawrilow and Joswig [GJ00].

Interesting discoveries about 3-dimensional convex polytopes were already made in ancient Greece. The treatise by Schläfli [Sch01] written in 1850–52 is usually mentioned as the beginning of modern theory, and several books were published around the turn of the century. We refer to Grünbaum [Grü67], Schrijver [Sch86], and to the other sources mentioned above for historical accounts.

The permutahedron mentioned in the introduction to this chapter was considered by Schoute [Sch11], and it arises by at least two other quite different and natural constructions (see [Zie94]).

There are several ways of proving the equivalence of H -polytopes and V -polytopes. Ours is inspired by a proof by Edmonds, as presented in Fukuda’s lecture notes (ETH Zürich). A classical algorithmic proof is provided by the *Fourier–Motzkin elimination procedure*, which proceeds by projections on coordinate hyperplanes; see [Zie94] for a detailed exposition. The *double-description method* is a similar algorithm formulated in the dual setting, and it is still one of the most efficient known computational methods. We will say a little more about the algorithmic problem of expressing the convex hull of a finite set as the intersection of half-spaces in the notes to Section 5.5.

One may ask, What is a “vertex description” of an *unbounded H -polyhedron*? Of course, it is not the convex hull of a finite set, but it can be expressed as the Minkowski sum $P + C$, where P is a V -polytope and C is a convex cone described as the convex hull of finitely many rays emanating from 0.

Exercises

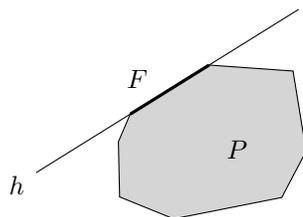
1. Verify that a d -dimensional simplex in \mathbf{R}^d can be expressed as the intersection of $d+1$ half-spaces. [2]
2. (a) Show that every convex polytope in \mathbf{R}^d is an orthogonal projection of a simplex of a sufficiently large dimension onto the space \mathbf{R}^d (which we consider embedded as a d -flat in some \mathbf{R}^n). [3]
 (b) Prove that every convex polytope P symmetric about 0 (i.e., with $P = -P$) is the affine image of a crosspolytope of a sufficiently large dimension. [3]

5.3 Faces of a Convex Polytope

The surface of the 3-dimensional cube consists of 8 “corner” points called vertices, 12 edges, and 6 squares called *facets*. According to the perhaps more usual terminology in 3-dimensional geometry, the facets would be called faces. But in the theory of convex polytopes, the word face has a slightly different meaning, defined below. For the cube, not only the squares but also the vertices and the edges are all called *faces* of the cube.

5.3.1 Definition (Face). A face of a convex polytope P is defined as

- either P itself, or
- a subset of P of the form $P \cap h$, where h is a hyperplane such that P is fully contained in one of the closed half-spaces determined by h .



We observe that each face of P is a convex polytope. This is because P is the intersection of finitely many half-spaces and h is the intersection of two half-spaces, so the face is an H -polyhedron, and moreover, it is bounded.

If P is a polytope of dimension d , then its faces have dimensions $-1, 0, 1, \dots, d$, where -1 is, by definition, the dimension of the empty set. A face of dimension j is also called a j -face.

Names of faces. The 0-faces are called *vertices*, the 1-faces are called *edges*, and the $(d-1)$ -faces of a d -dimensional polytope are called *facets*. The $(d-2)$ -faces of a d -dimensional polytope are *ridges*; in the familiar 3-dimensional situation, edges = ridges. For example, the 3-dimensional cube has 28 faces in total: the empty face, 8 vertices, 12 edges, 6 facets, and the whole cube.

The following proposition shows that each V -polytope is the convex hull of its vertices, and that the faces can be described combinatorially: They are the convex hulls of certain subsets of vertices. This includes some intuitive facts such as that each edge connects two vertices.

A helpful notion is that of an *extremal point* of a set: For a set $X \subseteq \mathbf{R}^d$, a point $x \in X$ is extremal if $x \notin \text{conv}(X \setminus \{x\})$.

5.3.2 Proposition. *Let $P \subset \mathbf{R}^d$ be a (bounded) convex polytope.*

- (i) (“Vertices are extremal”) *The extremal points of P are exactly its vertices, and P is the convex hull of its vertices.*
- (ii) (“Face of a face is a face”) *Let F be a face of P . The vertices of F are exactly those vertices of P that lie in F . More generally, the faces of F are exactly those faces of P that are contained in F .*

The proof is not essential for our further considerations, and it is given at the end of this section (but Exercise 9 below illustrates that things are not quite as simple as it might perhaps seem). The proposition has an appropriate analogue for polyhedra, but in order to avoid technicalities, we treat the bounded case only.

Examples. A d -dimensional simplex has been defined as the convex hull of a $(d+1)$ -point affinely independent set V . It is easy to see that each subset of V determines a face of the simplex. Thus, there are $\binom{d+1}{k+1}$ faces of dimension k , $k = -1, 0, \dots, d$, and 2^{d+1} faces in total.

The d -dimensional crosspolytope has $V = \{e_1, -e_1, \dots, e_d, -e_d\}$ as the vertex set. A proper subset $F \subset V$ determines a face if and only if there is

no i such that both $e_i \in F$ and $-e_i \in F$ (Exercise 2). It follows that there are 3^d+1 faces, including the empty one and the whole crosspolytope.

The nonempty faces of the d -dimensional cube $[-1, 1]^d$ correspond to vectors $v \in \{-1, 1, 0\}^d$. The face corresponding to such v has the vertex set $\{u \in \{-1, 1\}^d: u_i = v_i \text{ for all } i \text{ with } v_i \neq 0\}$. Geometrically, the vector v is the center of gravity of its face.

The face lattice. Let $\mathcal{F}(P)$ be the set of all faces of a (bounded) convex polytope P (including the empty face \emptyset of dimension -1). We consider the partial ordering of $\mathcal{F}(P)$ by inclusion.

5.3.3 Definition (Combinatorial equivalence). Two convex polytopes P and Q are called combinatorially equivalent if $\mathcal{F}(P)$ and $\mathcal{F}(Q)$ are isomorphic as partially ordered sets.

We are going to state some properties of the partially ordered set $\mathcal{F}(P)$ without proofs. These are not difficult and can be found in [Zie94].

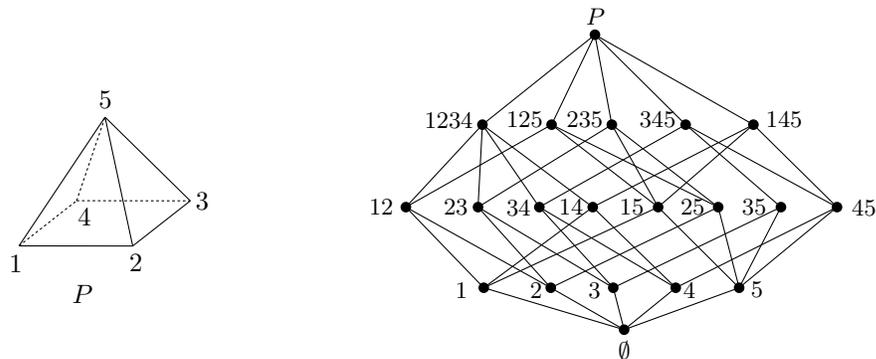
It turns out that $\mathcal{F}(P)$ is a lattice (a partially ordered set satisfying additional axioms). We recall that this means the following two conditions:

- *Meets condition:* For any two faces $F, G \in \mathcal{F}(P)$, there exists a face $M \in \mathcal{F}(P)$, called the *meet* of F and G , that is contained in both F and G and contains all other faces contained in both F and G .
- *Joins condition:* For any two faces $F, G \in \mathcal{F}(P)$, there exists a face $J \in \mathcal{F}(P)$, called the *join* of F and G , that contains both F and G and is contained in all other faces containing both F and G .

The meet of two faces is their geometric intersection $F \cap G$.

For verifying the joins and meets conditions, it may be helpful to know that for a finite partially ordered set possessing the minimum element and the maximum element, the meets condition is equivalent to the joins condition, and so it is enough to check only one of the conditions.

Here is the face lattice of a 3-dimensional pyramid:



The vertices are numbered 1–5, and the faces are labeled by the vertex sets.

The face lattice is *graded*, meaning that every maximal chain has the same length (the rank of a face F is $\dim(F)+1$). Quite obviously, it is *atomic*: Every face is the join of its vertices. A little less obviously, it is *coatomic*; that is, every face is the meet (intersection) of the facets containing it. An important consequence is that combinatorial type of a polytope is determined by the vertex–facet incidences. More precisely, if we know the dimension and all subsets of vertices that are vertex sets of facets (but without knowing the coordinates of the vertices, of course), we can uniquely reconstruct the whole face lattice in a simple and purely combinatorial way.

Face lattices of convex polytopes have several other nice properties, but no full algebraic characterization is known, and the problem of deciding whether a given lattice is a face lattice is algorithmically difficult (even for 4-dimensional polytopes).

The face lattice can be a suitable representation of a convex polytope in a computer. Each j -face is connected by pointers to its $(j-1)$ -faces and to the $(j+1)$ -faces containing it. On the other hand, it is a somewhat redundant representation: Recall that the vertex–facet incidences already contain the full information, and for some applications, even less data may be sufficient, say the graph of the polytope.

The dual polytope. Let P be a convex polytope containing the origin in its interior. Then the dual set P^* is also a polytope; we have verified this in the proof of Theorem 5.2.2.

5.3.4 Proposition. *For each $j = -1, 0, \dots, d$, the j -faces of P are in a bijective correspondence with the $(d-j-1)$ -faces of P^* . This correspondence also reverses inclusion; in particular, the face lattice of P^* arises by turning the face lattice of P upside down.*

Again we refer to the reader’s diligence or to [Zie94] for a proof. Let us examine a few examples instead.

Among the five regular Platonic solids, the cube and the octahedron are dual to each other, the dodecahedron and the icosahedron are also dual, and the tetrahedron is dual to itself. More generally, if we have a 3-dimensional convex polytope and G is its graph, then the graph of the dual polytope is the dual graph to G , in the usual graph-theoretic sense. The dual of a d -simplex is a d -simplex, and the d -dimensional cube and the d -dimensional crosspolytope are dual to each other.

We conclude with two notions of polytopes “in general position.”

5.3.5 Definition (Simple and simplicial polytopes). *A polytope P is called simplicial if each of its facets is a simplex (this happens, in particular, if the vertices of P are in general position, but general position is not necessary). A d -dimensional polytope P is called simple if each of its vertices is contained in exactly d facets.*

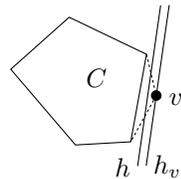
The faces of a simplex are again simplices, and so each proper face of a simplicial polytope is a simplex. Among the five Platonic solids, the tetrahedron, the octahedron, and the icosahedron are simplicial; and the tetrahedron, the cube, and the dodecahedron are simple. Crosspolytopes are simplicial, and cubes are simple. An example of a polytope that is neither simplicial nor simple is the 4-sided pyramid used in the illustration of the face lattice.

The dual of a simple polytope is simplicial, and vice versa. For a simple d -dimensional polytope, a small neighborhood of each vertex looks combinatorially like a neighborhood of a vertex of the d -dimensional cube. Thus, for each vertex v of a d -dimensional simple polytope, there are d edges emanating from v , and each k -tuple of these edges uniquely determines one k -face incident to v . Consequently, v belongs to $\binom{d}{k}$ k -faces, $k = 0, 1, \dots, d$.

Proof of Proposition 5.3.2. In (i) (“vertices are extremal”), we assume that P is the convex hull of a finite point set. Among all such sets, we fix one that is inclusion-minimal and call it V_0 . Let V_v be the vertex set of P , and let V_e be the set of all extremal points of P . We prove that $V_0 = V_v = V_e$, which gives (i). We have $V_e \subseteq V_0$ by the definition of an extremal point.

Next, we show that $V_v \subseteq V_e$. If $v \in V_v$ is a vertex of P , then there is a hyperplane h with $P \cap h = \{v\}$, and all of $P \setminus \{v\}$ lies in one of the open half-spaces defined by h . Hence $P \setminus \{v\}$ is convex, which means that v is an extremal point of P , and so $V_v \subseteq V_e$.

Finally we verify $V_0 \subseteq V_v$. Let $v \in V_0$; by the inclusion-minimality of V_0 , we get that $v \notin C = \text{conv}(V_0 \setminus \{v\})$. Since C and $\{v\}$ are disjoint compact convex sets, they can be strictly separated by a hyperplane h . Let h_v be the hyperplane parallel to h and containing v ; this h_v has all points of $V_0 \setminus \{v\}$ on one side.

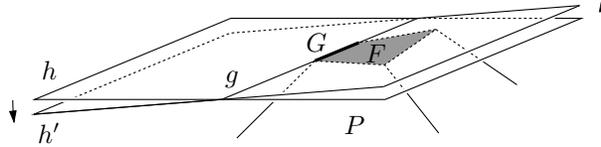


We want to show that $P \cap h_v = \{v\}$ (then v is a vertex of P , and we are done). The set $P \setminus h_v = \text{conv}(V_0) \setminus h_v$, being the intersection of a convex set with an open half-space, is convex. Any segment vx , where $x \in P \setminus h_v$, shares only the point v with the hyperplane h_v , and so $(P \setminus h_v) \cup \{v\}$ is convex as well. Since this set contains V_0 and is convex, it contains $P = \text{conv}(V_0)$, and so $P \cap h_v = \{v\}$ indeed.

As for (ii) (“face of a face is a face”), it is clear that a face G of P contained in F is a face of F too (use the same witnessing hyperplane). For the reverse direction, we begin with the case of vertices. By a consideration similar to that at the end of the proof of (i), we see that $F = \text{conv}(V) \cap h = \text{conv}(V \cap h)$,

where h is a hyperplane defining F . Hence all the extremal points of F , which by (i) are exactly the vertices of F , are in V .

Finally, let F be a face of P defined by a hyperplane h , and let $G \subset F$ be a face of F defined by a hyperplane g within h ; that is, g is a $(d-2)$ -dimensional affine subspace of h with $G = g \cap F$ and with all of F on one side. Let γ be the closed half-space bounded by h with $P \subset \gamma$. We start rotating the boundary h of γ around g in the direction such that the rotated half-space γ' still contains F .



If we rotate by a sufficiently small amount, then all the vertices of P not lying in F are still in the interior of γ' . At the same time, the interior of γ' contains all the vertices of F not lying in G , while all the vertices of G remain on the boundary h' of γ' . So h' defines a face of P (since all of P is on one side), and this face has the same vertex set as G , and so it equals G by the first part of (ii) proved above. \square

Exercises

- Verify that if $V \subset \mathbf{R}^d$ is affinely independent, then each subset $F \subseteq V$ determines a face of the simplex $\text{conv}(V)$. \square
- Verify the description of the faces of the cube and of the crosspolytope given in the text. \square
- Consider the $(n-1)$ -dimensional permutahedron as defined in the introduction to this chapter.
 - Verify that it really has $n!$ vertices corresponding to the permutations of $\{1, 2, \dots, n\}$. \square
 - Describe all faces of the permutahedron combinatorially (what sets of permutations are vertex sets of faces?). \square
 - Determine the dimensions of the faces found in (b). In particular, show that the facets correspond to ordered partitions (A, B) of $\{1, 2, \dots, n\}$, $A, B \neq \emptyset$, and count them. \square
- Let $P \subset \mathbf{R}^4 = \text{conv}\{\pm e_i \pm e_j : i, j = 1, 2, 3, 4, i \neq j\}$, where e_1, \dots, e_4 is the standard basis (this P is called the *24-cell*). Describe the face lattice of P and prove that P is combinatorially equivalent to P^* (in fact, P can be obtained from P^* by an isometry and scaling). \square
- Using Proposition 5.3.2, prove the following:
 - If F is a face of a convex polytope P , then F is the intersection of P with the affine hull of F . \square
 - If F and G are faces of a convex polytope P , then $F \cap G$ is a face, too. \square

6. Let P be a convex polytope in \mathbf{R}^3 containing the origin as an interior point, and let F be a j -face of P , $j = 0, 1, 2$.
 - (a) Give a precise definition of the face F' of the dual polytope P^* corresponding to F (i.e., describe F' as a subset of \mathbf{R}^3). \square
 - (b) Verify that F' is indeed a face of P^* . \square
7. Let $V \subset \mathbf{R}^d$ be the vertex set of a convex polytope and let $U \subset V$. Prove that U is the vertex set of a face of $\text{conv}(V)$ if and only if the affine hull of U is disjoint from $\text{conv}(V \setminus U)$. \square
8. Prove that the graph of any 3-dimensional convex polytope is 3-connected; i.e., removing any 2 vertices leaves the graph connected. \square
9. Let C be a convex set. Call a point $x \in C$ *exposed* if there is a hyperplane h with $C \cap h = \{x\}$ and all the rest of C on one side. For convex polytopes, exposed points are exactly the vertices, and we have shown that any extremal point is also exposed. Find an example of a compact convex set $C \subset \mathbf{R}^2$ with an extremal point that is not exposed. \square
10. (On extremal points) For a set $X \subseteq \mathbf{R}^d$, let $\text{ex}(X) = \{x \in X : x \notin \text{conv}(X \setminus \{x\})\}$ denote the set of extremal points of X .
 - (a) Find a convex set $C \subseteq \mathbf{R}^d$ with $C \neq \text{conv}(\text{ex}(C))$. \square
 - (b) Find a compact convex $C \subseteq \mathbf{R}^3$ for which $\text{ex}(C)$ is not closed. \square
 - (c) By modifying the proof of Proposition 5.3.2, prove that $C = \text{conv}(\text{ex}(C))$ for every compact convex $C \subset \mathbf{R}^d$ (this is a finite-dimensional version of the well known *Krein–Milman theorem*). \square

5.4 Many Faces: The Cyclic Polytopes

A convex polytope P can be given to us by the list of vertices. How difficult is it to recover the full face lattice, or, more modestly, a representation of P as an intersection of half-spaces? The first question to ask is how large the face lattice or the collection of half-spaces can be, compared to the number of vertices. That is, what is the maximum total number of faces, or the maximum number of facets, of a convex polytope in \mathbf{R}^d with n vertices? The dual question is, of course, the maximum number of faces or vertices of a bounded intersection of n half-spaces in \mathbf{R}^d .

Let $f_j = f_j(P)$ denote the number of j -faces of a polytope P . The vector (f_0, f_1, \dots, f_d) is called the *f-vector* of P . We thus assume $f_0 = n$ and we are interested in estimating the maximum value of f_{d-1} and of $\sum_{k=0}^d f_k$.

In dimensions 2 and 3, the situation is simple and favorable. For $d = 2$, our polytope is a convex polygon with n vertices and n edges, and so $f_0 = f_1 = n$, $f_2 = 1$. The *f-vector* is even determined uniquely.

A 3-dimensional polytope can be regarded as a drawing of a planar graph, in our case with n vertices. By well-known results for planar graphs, we have $f_1 \leq 3n - 6$ and $f_2 \leq 2n - 4$. Equalities hold if and only if the polytope is simplicial (all facets are triangles).

In both cases the total number of faces is linear in n . But as the dimension grows, polytopes become much more complicated. First of all, even the total number of faces of the most innocent convex polytope, the d -dimensional simplex, is *exponential* in d . But here we consider d fixed and relatively small, and we investigate the dependence on the number of vertices n .

Still, as we will see, for every $n \geq 5$ there is a 4-dimensional convex polytope with n vertices and with every two vertices connected by an edge, i.e., with $\binom{n}{2}$ edges! This looks counterintuitive, but our intuition is based on the 3-dimensional case. In any fixed dimension d , the number of facets can be of order $n^{\lfloor d/2 \rfloor}$, which is rather disappointing for someone wishing to handle convex polytopes efficiently. On the other hand, complete desperation is perhaps not appropriate: Certainly not all polytopes exhibit this very bad behavior. For example, it is known that if we choose n points uniformly at random in the unit ball B^d , then the expected number of faces of their convex hull is only $o(n)$, for every fixed d .

It turns out that the number of faces for a given dimension and number of vertices is the largest possible for so-called *cyclic polytopes*, to be introduced next. First we define a very useful curve in \mathbf{R}^d .

5.4.1 Definition (Moment curve). *The curve $\gamma = \{(t, t^2, \dots, t^d) : t \in \mathbf{R}\}$ in \mathbf{R}^d is called the moment curve.*

5.4.2 Lemma. *Any hyperplane h intersects the moment curve γ in at most d points. If there are d intersections, then h cannot be tangent to γ , and thus at each intersection, γ passes from one side of h to the other.*

Proof. A hyperplane h can be expressed by the equation $\langle a, x \rangle = b$, or in coordinates $a_1x_1 + a_2x_2 + \dots + a_dx_d = b$. A point of γ has the form (t, t^2, \dots, t^d) , and if it lies in h , we obtain $a_1t + a_2t^2 + \dots + a_dt^d - b = 0$. This means that t is a root of a nonzero polynomial $p_h(t)$ of degree at most d , and hence the number of intersections of h with γ is at most d . If there are d distinct roots, then they must be all simple. At a simple root, the polynomial $p_h(t)$ changes sign, and this means that the curve γ passes from one side of h to the other. \square

As a corollary, we see that every d points of the moment curve are affinely independent, for otherwise, we could pass a hyperplane through them plus one more point of γ . So the moment curve readily supplies *explicit* examples of point sets in general position.

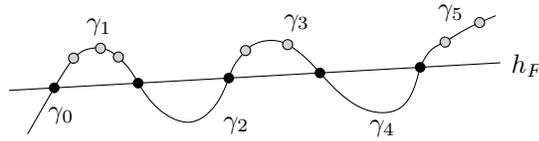
5.4.3 Definition (Cyclic polytope). *The convex hull of finitely many points on the moment curve is called a cyclic polytope.*

How many facets does a cyclic polytope have? Each facet is determined by a d -tuple of vertices, and distinct d -tuples determine distinct facets. Here is a criterion telling us exactly which d -tuples determine facets.

5.4.4 Proposition (Gale's evenness criterion). Let V be the vertex set of a cyclic polytope P considered with the linear ordering \leq along the moment curve (larger vertices have larger values of the parameter t). Let $F = \{v_1, v_2, \dots, v_d\} \subseteq V$ be a d -tuple of vertices of P , where $v_1 < v_2 < \dots < v_d$. Then F determines a facet of P if and only if for any two vertices $u, v \in V \setminus F$, the number of vertices $v_i \in F$ with $u < v_i < v$ is even.

Proof. Let h_F be the hyperplane affinely spanned by F . Then F determines a facet if and only if all the points of $V \setminus F$ lie on the same side of h_F .

Since the moment curve γ intersects h_F in exactly d points, namely at the points of F , it is partitioned into $d+1$ pieces, say $\gamma_0, \dots, \gamma_d$, each lying completely in one of the half-spaces, as is indicated in the drawing:



Hence, if the vertices of $V \setminus F$ are all contained in the odd-numbered pieces $\gamma_1, \gamma_3, \dots$, as in the picture, or if they are all contained in the even-numbered pieces $\gamma_0, \gamma_2, \dots$, then F determines a facet. This condition is equivalent to Gale's criterion. \square

Now we can count the facets.

5.4.5 Theorem. The number of facets of a d -dimensional cyclic polytope with n vertices ($n \geq d+1$) is

$$\binom{n - \lfloor d/2 \rfloor}{\lfloor d/2 \rfloor} + \binom{n - \lfloor d/2 \rfloor - 1}{\lfloor d/2 \rfloor - 1} \text{ for } d \text{ even, and}$$

$$2 \binom{n - \lfloor d/2 \rfloor - 1}{\lfloor d/2 \rfloor} \text{ for } d \text{ odd.}$$

For fixed d , this has the order of magnitude $n^{\lfloor d/2 \rfloor}$.

Proof. The number of facets equals the number of ways of placing d black circles and $n - d$ white circles in a row in such a way that we have an even number of black circles between each two white circles.

Let us say that an arrangement of black and white circles is *paired* if any contiguous segment of black circles has an even length (the arrangements permitted by Gale's criterion need not be paired because of the initial and final segments). The number of paired arrangements of $2k$ black circles and $n - 2k$ white circles is $\binom{n-k}{k}$, since by deleting every second black circle we get a one-to-one correspondence with selections of the positions of k black circles among $n - k$ possible positions.

Let us return to the original problem, and first consider an *odd* $d = 2k+1$. In a valid arrangement of circles, we must have an odd number of consecutive black circles at the beginning or at the end (but not both). In the former case, we delete the initial black circle, and we get a paired arrangement of $2k$ black and $n-1-2k$ white circles. In the latter case, we similarly delete the black circle at the end and again get a paired arrangement as in the first case. This establishes the formula in the theorem for odd d .

For *even* $d = 2k$, the number of initial consecutive black circles is either odd or even. In the even case, we have a paired arrangement, which contributes $\binom{n-k}{k}$ possibilities. In the odd case, we also have an odd number of consecutive black circles at the end, and so by deleting the first and last black circles we obtain a paired arrangement of $2(k-1)$ black circles and $n-2k$ white circles. This contributes $\binom{n-k-2}{k-1}$ possibilities. \square

Bibliography and remarks. The convex hull of the moment curve was studied by by Carathéodory [Car07]. In the 1950s, Gale constructed neighborly polytopes by induction. Cyclic polytopes and the evenness criterion appear in Gale [Gal63]. The moment curve is an important object in many other branches besides the theory of convex polytopes. For example, in elementary algebraic topology it is used for proving that every (at most countable) d -dimensional simplicial complex has a geometric realization in \mathbf{R}^{2d+1} .

Exercises

- (a) Show that if V is a finite subset of the moment curve, then all the points of V are extreme in $\text{conv}(V)$; that is, they are vertices of the corresponding cyclic polytope. \square
 (b) Show that any two cyclic polytopes in \mathbf{R}^d with n vertices are combinatorially the same: They have isomorphic face lattices. Thus, we can speak of *the* cyclic polytope. \square
- Show that for cyclic polytopes in dimensions 4 and higher, every pair of vertices is connected by an edge. For dimension 4 and two arbitrary vertices, write out explicitly the equation of a hyperplane intersecting the cyclic polytope exactly in this edge. \square
- Determine the f -vector of a cyclic polytope with n vertices in dimensions 4, 5, and 6. \square

5.5 The Upper Bound Theorem

The upper bound theorem, one of the earlier major achievements of the theory of convex polytopes, claims that the cyclic polytope has the largest possible number of faces.

5.5.1 Theorem (Upper bound theorem). *Among all d -dimensional convex polytopes with n vertices, the cyclic polytope maximizes the number of faces of each dimension.*

In this section we prove only an approximate result, which gives the correct order of magnitude for the maximum number of facets.

5.5.2 Proposition (Asymptotic upper bound theorem). *A d -dimensional convex polytope with n vertices has at most $2\binom{n}{\lfloor d/2 \rfloor}$ facets and no more than $2^{d+1}\binom{n}{\lfloor d/2 \rfloor}$ faces in total. For d fixed, both quantities thus have the order of magnitude $n^{\lfloor d/2 \rfloor}$.*

First we establish this proposition for simplicial polytopes, in the following form.

5.5.3 Proposition. *Let P be a d -dimensional simplicial polytope. Then*

- (a) $f_0(P) + f_1(P) + \cdots + f_d(P) \leq 2^d f_{d-1}(P)$, and
- (b) $f_{d-1}(P) \leq 2f_{\lfloor d/2 \rfloor - 1}(P)$.

This implies Proposition 5.5.2 for simplicial polytopes, since the number of $(\lfloor d/2 \rfloor - 1)$ -faces is certainly no bigger than $\binom{n}{\lfloor d/2 \rfloor}$, the number of all $\lfloor d/2 \rfloor$ -tuples of vertices.

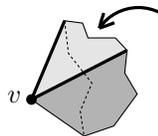
Proof of Proposition 5.5.3. We pass to the dual polytope P^* , which is simple. Now we need to prove $\sum_{k=0}^d f_k(P^*) \leq 2^d f_0(P^*)$ and $f_0(P^*) \leq 2f_{\lfloor d/2 \rfloor}(P^*)$.

Each face of P^* has at least one vertex, and every vertex of a simple d -polytope is incident to 2^d faces, which gives the first inequality.

We now bound the number of vertices in terms of the number of $\lfloor d/2 \rfloor$ -faces. This is the heart of the proof, and it shows where the mysterious exponent $\lfloor d/2 \rfloor$ comes from.

Let us rotate the polytope P^* so that no two vertices share the x_d -coordinate (i.e., no two vertices have the same vertical level).

Consider a vertex v with the d edges emanating from it. By the pigeonhole principle, there are at least $\lfloor d/2 \rfloor$ edges directed upwards or at least $\lfloor d/2 \rfloor$ edges directed downwards. In the former case, every $\lfloor d/2 \rfloor$ -tuple of edges going up determines a $\lfloor d/2 \rfloor$ -face for which v is the lowest vertex. In the latter case, every $\lfloor d/2 \rfloor$ -tuple of edges going down determines a $\lfloor d/2 \rfloor$ -face for which v is the highest vertex. Here is an illustration, unfortunately for the not too interesting 3-dimensional case, showing a situation with 2 edges going up and the corresponding 2-dimensional face having v as the lowest vertex:



We have exhibited at least one $\lceil d/2 \rceil$ -face for which v is the lowest vertex or the highest vertex. Since the lowest vertex and the highest vertex are unique for each face, the number of vertices is no more than twice the number of $\lceil d/2 \rceil$ -faces. \square

Warning. For simple polytopes, the total combinatorial complexity is proportional to the number of vertices, and for simplicial polytopes it is proportional to the number of facets (considering the dimension fixed, that is). For polytopes that are neither simple nor simplicial, the number of faces of intermediate dimensions can have larger order of magnitude than both the number of facets and the number of vertices; see Exercise 1.

Nonsimplicial polytopes. To prove the asymptotic upper bound theorem, it remains to deal with nonsimplicial polytopes. This is done by a perturbation argument, similar to numerous other results where general position is convenient for the proof but where we want to show that the result holds in degenerate cases as well. In most instances in this book, the details of perturbation arguments are omitted, but here we make an exception, since the proof seems somewhat nontrivial.

5.5.4 Lemma. *For any d -dimensional convex polytope P there exists a d -dimensional simplicial polytope Q with $f_0(P) = f_0(Q)$ and $f_k(Q) \geq f_k(P)$ for all $k = 1, 2, \dots, d$.*

Proof. The basic idea is very simple: Move (perturb) every vertex of P by a very small amount, in such a way that the vertices are in general position, and show that each k -face of P gives rise to at least one k -face of the perturbed polytope. There are several ways of doing this proof.

We process the vertices one by one. Let V be the vertex set of P and let $v \in V$. The operation of ε -pushing v is as follows: We choose a point v' lying in the interior of P , at distance at most ε from v , and on no hyperplane determined by the points of V , and we set $V' = (V \setminus \{v\}) \cup \{v'\}$. If we successively ε_v -push each vertex v of the polytope, the resulting vertex set is in general position and we have a simple polytope.

It remains to show that for any polytope P with vertex set V and any $v \in V$, there is an $\varepsilon > 0$ such that ε -pushing v does not decrease the number of faces.

Let $U \subset V$ be the vertex set of a k -face of P , $0 \leq k \leq d-1$, and let V' arise from V by ε -pushing v . If $v \notin U$, then no doubt, U determines a face of $\text{conv}(V')$, and so we assume that $v \in U$. First suppose that v lies in the affine hull of $U \setminus \{v\}$; we claim that then $U \setminus \{v\}$ determines a k -face of $\text{conv}(V')$. This follows easily from the criterion in Exercise 5.3.7: A subset $U \subset V$ is the vertex set of a face of $\text{conv}(V)$ if and only if the affine hull of U is disjoint from $\text{conv}(V \setminus U)$. We leave a detailed argument to the reader (one must use the fact that v is pushed inside).

If v lies outside of the affine hull of $U \setminus \{v\}$, then we want to show that $U' = (U \setminus \{v\}) \cup \{v'\}$ determines a k -face of $\text{conv}(V')$. The affine hull of U is disjoint from the compact set $\text{conv}(V \setminus U)$. If we move v continuously by a sufficiently small amount, the affine hull of U moves continuously, and so there is an $\varepsilon > 0$ such that if we move v within ε from its original position, the considered affine hull and $\text{conv}(V \setminus U)$ remain disjoint. \square

Bibliography and remarks. The upper bound theorem was conjectured by Motzkin in 1957 and proved by McMullen [McM70]. Many partial results have been obtained in the meantime. Perhaps most notably, Klee [Kle64] found a simple proof for polytopes with not too few vertices (at least about d^2 vertices in dimension d). That proof applies to simplicial complexes much more general than the boundary complexes of simplicial polytopes: It works for Eulerian pseudomanifolds and, in particular, for all *simplicial spheres*, i.e., simplicial complexes homeomorphic to S^{d-1} . Presentations of McMullen's proof and Klee's proof can be found in Ziegler's book [Zie94]. A nice variation was described by Alon and Kalai [AK85].

The proof of Lemma 5.5.4 by pushing vertices inside is similar to an argument in Klee [Kle64], but he proves more and presents the proof in more detail.

Convex hull computation. What does it mean to compute the convex hull of a given n -point set $V \subset \mathbf{R}^d$? One possible answer, briefly touched upon in the notes to Section 5.2, is to express $\text{conv}(V)$ as the intersection of half-spaces and to compute the vertex sets of all facets. (As we know, the face lattice can be reconstructed from this information purely combinatorially; see Kaibel and Pfetsch [KP01] for an efficient algorithm.) Of course, for some applications it may be sufficient to know much less about the convex hull, say only the graph of the polytope or only the list of its vertices, but here we will discuss only algorithms for computing all the vertex–facet incidences or the whole face lattice. For a more detailed overview of convex hull algorithms see, e.g., Seidel [Sei97].

For the dimension d considered fixed, there is a quite simple and practical randomized algorithm. It computes the convex hull of n points in \mathbf{R}^d in expected time $O(n^{\lfloor d/2 \rfloor} + n \log n)$ (Seidel [Sei91], simplifying Clarkson and Shor [CS89]), and also a very complicated but deterministic algorithm with the same asymptotic running time (Chazelle [Cha93b]; somewhat simplified in Brönnimann, Chazelle, and Matoušek [BCM99]). This is worst-case optimal, since an n -vertex polytope may have about $n^{\lfloor d/2 \rfloor}$ facets. There are also output-sensitive algorithms, whose running time depends on the total number f of faces of the resulting polytope. Recent results in this direction, including an algorithm that computes the convex hull of n points in general posi-

tion in \mathbf{R}^d (d fixed) in time $O(n \log f + (nf)^{1-1/(\lfloor d/2 \rfloor + 1)} (\log n)^{c(d)})$, can be found in Chan [Cha00b].

Still, none of the known algorithms is theoretically fully satisfactory, and practical computation of convex hulls even in moderate dimensions, say 10 or 20, can be quite challenging. Some of the algorithms are too complicated and with too large constants hidden in the asymptotic notation to be of practical value. Algorithms requiring general position of the points are problematic for highly degenerate point configurations (which appear in many applications), since small perturbations used to achieve general position often increase the number of faces tremendously. Some of the randomized algorithms compute intermediate polytopes that can have many more faces than the final result. Often we are interested just in the vertex–facet incidences, but many algorithms compute all faces, whose number can be much larger, or even a triangulation of every face, which may again increase the complexity. Such problems of existing algorithms are discussed in Avis, Bremner, and Seidel [ABS97].

For actual computations, simple and theoretically suboptimal algorithms are often preferable. One of them is the double-description method mentioned earlier, and another algorithm that seems to behave well in many difficult instances is the *reverse search* of Avis and Fukuda [AF92]. It enumerates the vertices of the intersection of a given set H of half-spaces one by one, using quite small storage. Conceptually, one thinks of optimizing a generic linear function over $\bigcap H$ by a simplex algorithm with Bland’s rule. This defines a spanning tree in the graph of the polytope, and this tree is searched depth-first starting from the optimum vertex, essentially by running the simplex algorithm “backwards.” The main problem of this algorithm is with degenerate vertices of high degree, which may correspond to an enormous number of bases in the simplex algorithm.

Exercises

1. (a) Let P be a k -dimensional convex polytope in \mathbf{R}^k , and Q an ℓ -dimensional convex polytope in \mathbf{R}^ℓ . Show that the Cartesian product $P \times Q \subset \mathbf{R}^{k+\ell}$ is a convex polytope of dimension $k + \ell$. □
- (b) If F is an i -face of P , and G is a j -face of Q , $i, j \geq 0$, then $F \times G$ is an $(i + j)$ -face of $P \times Q$. Moreover, this yields all the nonempty faces of $P \times Q$. □
- (c) Using the product of suitable polytopes, find an example of a “fat-lattice” polytope, i.e., a polytope for which the total number of faces has a larger order of magnitude than the number of vertices plus the number of facets together (the dimension should be a constant). □

(d) Show that the following yields a 5-dimensional fat-lattice polytope: The convex hull of two regular n -gons whose affine hulls are skew 2-flats in \mathbf{R}^5 . \square

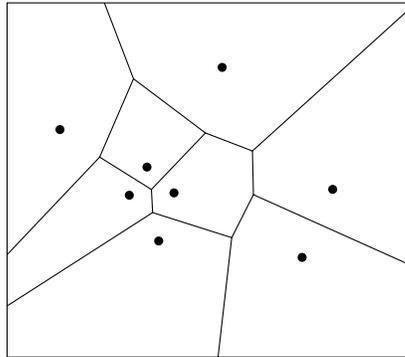
For recent results on fat-lattice polytopes see Eppstein, Kuperberg, and Ziegler [EKZ03].

5.7 Voronoi Diagrams

Consider a finite set $P \subset \mathbf{R}^d$. For each point $p \in P$, we define a region $reg(p)$, which is the “sphere of influence” of the point p : It consists of the points $x \in \mathbf{R}^d$ for which p is the closest point among the points of P . Formally,

$$reg(p) = \{x \in \mathbf{R}^d: \text{dist}(x, p) \leq \text{dist}(x, q) \text{ for all } q \in P\},$$

where $\text{dist}(x, y)$ denotes the Euclidean distance of the points x and y . The *Voronoi diagram* of P is the set of all regions $reg(p)$ for $p \in P$. (More precisely, it is the cell complex induced by these regions; that is, every intersection of a subset of the regions is a face of the Voronoi diagram.) Here an example of the Voronoi diagram of a point set in the plane:



(Of course, the Voronoi diagram is clipped by a rectangle so that it fits into a finite page.) The points of P are traditionally called the *sites* in the context of Voronoi diagrams.

5.7.1 Observation. *Each region $reg(p)$ is a convex polyhedron with at most $|P|-1$ facets.*

Indeed,

$$reg(p) = \bigcap_{q \in P \setminus \{p\}} \{x: \text{dist}(x, p) \leq \text{dist}(x, q)\}$$

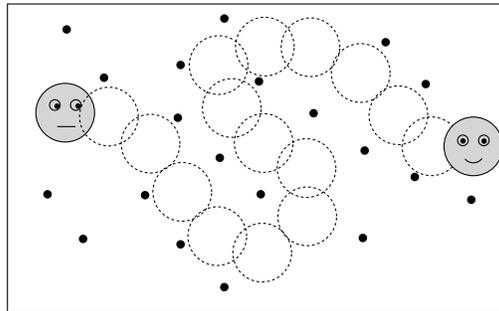
is an intersection of $|P| - 1$ half-spaces. \square

For $d = 2$, a Voronoi diagram of n points is a subdivision of the plane into n convex polygons (some of them are unbounded). It can be regarded as a drawing of a planar graph (with one vertex at the infinity, say), and hence it has a linear combinatorial complexity: n regions, $O(n)$ vertices, and $O(n)$ edges.

In the literature the Voronoi diagram also appears under various other names, such as the *Dirichlet tessellation*.

Examples of applications. Voronoi diagrams have been reinvented and used in various branches of science. Sometimes the connections are surprising. For instance, in archaeology, Voronoi diagrams help study cultural influences. Here we mention a few applications, mostly algorithmic.

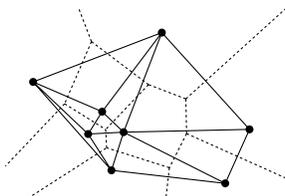
- (“Post office problem” or nearest neighbor searching) Given a point set P in the plane, we want to construct a data structure that finds the point of P nearest to a given query point x as quickly as possible. This problem arises directly in some practical situations or, more significantly, as a subroutine in more complicated problems. The query can be answered by determining the region of the Voronoi diagram of P containing x . For this problem (point location in a subdivision of the plane), efficient data structures are known; see, e.g., the book [dBvKOS97] or other introductory texts on computational geometry.
- (Robot motion planning) Consider a disk-shaped robot in the plane. It should pass among a set P of point obstacles, getting from a given start position to a given target position and touching none of the obstacles.



If such a passage is possible at all, the robot can always walk along the edges of the Voronoi diagram of P , except for the initial and final segments of the tour. This allows one to reduce the robot motion problem to a graph search problem: We define a subgraph of the Voronoi diagram consisting of the edges that are passable for the robot.

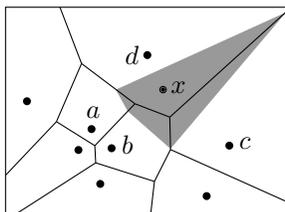
- (A nice triangulation: the Delaunay triangulation) Let $P \subset \mathbf{R}^2$ be a finite point set. In many applications one needs to construct a triangulation of P (that is, to subdivide $\text{conv}(P)$ into triangles with vertices at the points of P) in such a way that the triangles are not too skinny. Of course, for

some sets, some skinny triangles are necessary, but we want to avoid them as much as possible. One particular triangulation that is usually very good, and provably optimal with respect to several natural criteria, is obtained as the dual graph to the Voronoi diagram of P . Two points of P are connected by an edge if and only if their Voronoi regions share an edge.



If no 4 points of P lie on a common circle then this indeed defines a triangulation, called the *Delaunay triangulation*² of P ; see Exercise 5. The definition extends to points sets in \mathbf{R}^d in a straightforward manner.

- (Interpolation) Suppose that $f: \mathbf{R}^2 \rightarrow \mathbf{R}$ is some smooth function whose values are known to us only at the points of a finite set $P \subset \mathbf{R}^2$. We would like to interpolate f over the whole polygon $\text{conv}(P)$. Of course, we cannot really tell what f looks like outside P , but still we want a reasonable interpolation rule that provides a nice smooth function with the given values at P . Multidimensional interpolation is an extensive semiempirical discipline, which we do not seriously consider here; we explain only one elegant method based on Voronoi diagrams. To compute the interpolated value at a point $x \in \text{conv}(P)$, we construct the Voronoi diagram of P , and we overlay it with the Voronoi diagram of $P \cup \{x\}$.



The region of the new point x cuts off portions of the regions of some of the old points. Let w_p be the area of the part of $\text{reg}(p)$ in the Voronoi diagram of P that belongs to $\text{reg}(x)$ after inserting x . The interpolated value $f(x)$ is

$$f(x) = \sum_{p \in P} \frac{w_p}{\sum_{q \in P} w_q} f(p).$$

An analogous method can be used in higher dimensions, too.

² Being a transcription from Russian, the spelling of Delaunay's name varies in the literature. For example, in crystallography literature he is usually spelled "Delone."

Relation of Voronoi diagrams to convex polyhedra. We now show that Voronoi diagrams in \mathbf{R}^d correspond to certain convex polyhedra in \mathbf{R}^{d+1} .

First we define the *unit paraboloid* in \mathbf{R}^{d+1} :

$$U = \{x \in \mathbf{R}^{d+1}: x_{d+1} = x_1^2 + x_2^2 + \cdots + x_d^2\}.$$

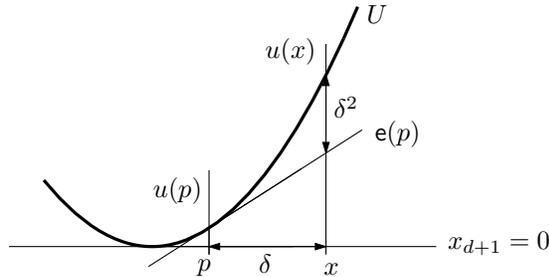
For $d = 1$, U is a parabola in the plane.

In the sequel, let us imagine the space \mathbf{R}^d as the hyperplane $x_{d+1} = 0$ in \mathbf{R}^{d+1} . For a point $p = (p_1, \dots, p_d) \in \mathbf{R}^d$, let $e(p)$ denote the hyperplane in \mathbf{R}^{d+1} with equation

$$x_{d+1} = 2p_1x_1 + 2p_2x_2 + \cdots + 2p_dx_d - p_1^2 - p_2^2 - \cdots - p_d^2.$$

Geometrically, $e(p)$ is the hyperplane tangent to the paraboloid U at the point $u(p) = (p_1, p_2, \dots, p_d, p_1^2 + \cdots + p_d^2)$ lying vertically above p . It is perhaps easier to remember this geometric definition of $e(p)$ and derive its equation by differentiation when needed. On the other hand, in the forthcoming proof we start out from the equation of $e(p)$, and as a by-product, we will see that $e(p)$ is the tangent to U at $u(p)$ as claimed.

5.7.2 Proposition. *Let $p, x \in \mathbf{R}^d$ be points and let $u(x)$ be the point of U vertically above x . Then $u(x)$ lies above the hyperplane $e(p)$ or on it, and the vertical distance of $u(x)$ to $e(p)$ is δ^2 , where $\delta = \text{dist}(x, p)$.*

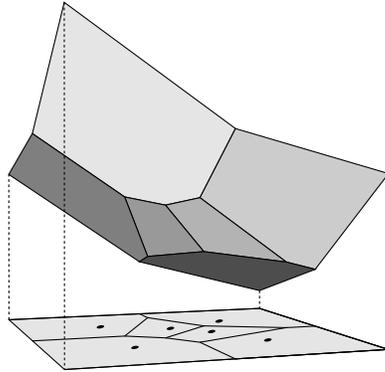


Proof. We just substitute into the equations of U and of $e(p)$. The x_{d+1} -coordinate of $u(x)$ is $x_1^2 + \cdots + x_d^2$, while the x_{d+1} -coordinate of the point of $e(p)$ above x is $2p_1x_1 + \cdots + 2p_dx_d - p_1^2 - \cdots - p_d^2$. The difference is $(x_1 - p_1)^2 + \cdots + (x_d - p_d)^2 = \delta^2$. \square

Let $\mathcal{E}(p)$ denote the half-space lying above the hyperplane $e(p)$. Consider an n -point set $P \subset \mathbf{R}^d$. By Proposition 5.7.2, $x \in \text{reg}(p)$ holds if and only if $e(p)$ is vertically closest to U at x among all $e(q)$, $q \in P$. Here is what we have derived:

5.7.3 Corollary. *The Voronoi diagram of P is the vertical projection of the facets of the polyhedron $\bigcap_{p \in P} \mathcal{E}(p)$ onto the hyperplane $x_{d+1} = 0$.* \square

Here is an illustration for a planar Voronoi diagram:



5.7.4 Corollary. *The maximum total number of faces of all regions of the Voronoi diagram of an n -point set in \mathbf{R}^d is $O(n^{\lceil d/2 \rceil})$.*

Proof. We know that the combinatorial complexity of the Voronoi diagram equals the combinatorial complexity of an H -polyhedron with at most n facets in \mathbf{R}^{d+1} . By intersecting this H -polyhedron with a large simplex we can obtain a bounded polytope with at most $n+d+2$ facets, and we have not decreased the number of faces compared to the original H -polyhedron. Then the dual version of the asymptotic upper bound theorem (Theorem 5.5.2) implies that the total number of faces is $O(n^{\lceil d/2 \rceil})$, since $\lfloor (d+1)/2 \rfloor = \lceil d/2 \rceil$. \square

The convex polyhedra in \mathbf{R}^{d+1} obtained from Voronoi diagrams in \mathbf{R}^d by the above construction are rather special, and so a lower bound for the combinatorial complexity of convex polytopes cannot be automatically transferred to Voronoi diagrams. But it turns out that the number of vertices of a Voronoi diagram on n points in \mathbf{R}^d can really be of order $n^{\lceil d/2 \rceil}$ (Exercise 2).

Let us remark that the trick used for transforming Voronoi diagrams to convex polyhedra is an example of a more general technique, called *linearization* or *Veronese mapping*, which will be discussed a little more in Section ???. This method sometimes allows us to convert a problem about algebraic curves or surfaces of bounded degree to a problem about k -flats in a suitable higher-dimensional space.

The farthest-point Voronoi diagram. The projection of the H -polyhedron $\bigcap_{p \in P} \mathcal{E}(p)^{\text{op}}$, where γ^{op} denotes the half-space opposite to γ , forms the *farthest-point Voronoi diagram*, in which each point $p \in P$ is assigned the regions of points for which it is the farthest point. It can be shown that all nonempty regions of this diagram are unbounded and they correspond precisely to the points appearing on the surface of $\text{conv}(P)$.

Bibliography and remarks. The concept of Voronoi diagrams independently emerged in various fields of science, for example as the *medial axis transform* in biology and physiology, the *Wigner–Seitz zones* in chemistry and physics, the *domains of action* in crystallography, and the *Thiessen polygons* in meteorology and geography. Apparently, the earliest documented reference to Voronoi diagrams is a picture in the famous *Principia Philosophiae* by Descartes from 1644 (that picture actually seems to show a power diagram, a generalization of the Voronoi diagram to sites with different strengths of influence). Mathematically, Voronoi diagrams were first introduced by Dirichlet [Dir50] and by Voronoi [Vor08] for the investigation of quadratic forms. For more information on the interesting history and a surprising variety of applications we refer to several surveys: Aurenhammer and Klein [AK00], Aurenhammer [Aur91], and the book Okabe, Boots, and Sugihara [OBS92]. Every computational geometry textbook also has at least a chapter devoted to Voronoi diagrams, and most papers on this subject appear in computational geometry.

The *Delaunay triangulation* (or, more correctly, the Delaunay tessellation, since it need not be a triangulation in general) was first considered by Voronoi as the dual to the Voronoi diagram, and later by Delaunay [Del34] with the definition given in Exercise 5(b) below. The Delaunay triangulation of a planar point set P optimizes several quality measures among all triangulations of P : It maximizes the minimum angle occurring in any triangle, minimizes the maximum circumradius of the triangles, maximizes the sum of inradii, and so on (see [AK00] for references). Such optimality properties can usually be proved by *local flipping*. We consider an arbitrary triangulation \mathcal{T} of a given finite $P \subset \mathbf{R}^2$ (say with no 4 cocircular points). If there is a 4-point $Q \subseteq P$ such that $\text{conv}(Q)$ is a quadrilateral triangulated by two triangles of \mathcal{T} but in such a way that these two triangles are not the Delaunay triangulation of Q , then the diagonal of Q can be flipped:



It can be shown that every sequence of such local flips is finite and finishes with the Delaunay triangulation of P (Exercise 7). This procedure has an analogue in higher dimensions, where it gives a simple and practically successful algorithm for computing Delaunay triangulations (and Voronoi diagrams); see, e.g., Edelsbrunner and Shah [ES96].

Generalizations of Voronoi diagrams. The example in the text with robot motion planning, as well as other applications, motivates various notions of generalized Voronoi diagrams. First, instead of the Euclidean distance, one can take various other distance functions, say the ℓ_p -metrics. Second, instead of the spheres of influence of points, we can consider the spheres of influence of other sites, such as disjoint polygons (this is what we get if we have a circular robot moving amidst polygonal obstacles). We do not attempt to survey the numerous results concerning such generalizations, again referring to [AK00]. Results on the combinatorial complexity of Voronoi diagrams under non-Euclidean metrics and/or for nonpoint sites will be mentioned in the notes to Section ??.

In another, very general, approach to Voronoi diagrams, one takes the Voronoi diagram induced by two objects as a primitive notion. So for every two objects we are given a partition of space into two regions separated by a *bisector*, and Voronoi diagrams for more than two objects are built using the 2-partitions for all pairs. If one postulates a few geometric properties of the bisectors, one gets a reasonable theory of Voronoi diagrams (the so-called *abstract Voronoi diagrams*), including efficient algorithms. So, for example, we do not even need a notion of distance at this level of generality. Abstract Voronoi diagrams (in the plane) were suggested by Klein [Kle89].

A geometrically significant generalization of the Euclidean Voronoi diagram is the *power diagram*: Each point $p \in P$ is assigned a real weight $w(p)$, and $\text{reg}(P) = \{x \in \mathbf{R}^d: \|x - p\|^2 - w(p) \leq \|x - q\|^2 - w(q) \text{ for all } q \in P\}$. While Voronoi diagrams in \mathbf{R}^d are projections of *certain* convex polyhedra in \mathbf{R}^{d+1} , the projection into \mathbf{R}^d of every intersection of finitely many nonvertical upper half-spaces in \mathbf{R}^{d+1} is a power diagram. Moreover, a hyperplane section of a power diagram is again a power diagram. Several other generalized Voronoi diagrams in \mathbf{R}^d (for example, with multiplicative weights of the sites) can be obtained by intersecting a suitable power diagram in \mathbf{R}^{d+1} with a simple surface and projecting into \mathbf{R}^d , which yields fast algorithms; see Aurenhammer and Imai [AI88].

Another generalization are *higher-order Voronoi diagrams*. The k th-order Voronoi diagram of a finite point set P assigns to each k -point $T \subseteq P$ the region $\text{reg}(T)$ consisting of all $x \in \mathbf{R}^d$ for which the points of T are the k nearest neighbors of x in P . The usual Voronoi diagram arises for $k = 1$, and the farthest-point Voronoi diagram for $k = |P| - 1$. The k th-order Voronoi diagram of $P \subset \mathbf{R}^d$ is the projection of the k th level facets in the arrangement of the hyperplanes $e(p)$, $p \in P$ (see Chapter 6 for these notions). Lee [Lee82] proved that the k th-order Voronoi diagram of n points in the plane has combinato-

rial complexity $O(k(n-k))$; this is better than the maximum possible complexity of level k in an arrangement of n arbitrary planes in \mathbf{R}^3 .

Applications of Voronoi diagrams are too numerous to be listed here, and we add only a few remarks to those already mentioned in the text. Using point location in Voronoi diagrams as in the post office problem, several basic computational problems in the plane can be solved efficiently, such as finding the *closest pair* in a point set or the *largest disk* contained in a given polygon and not containing any of the given points.

Besides providing good triangulations, the Delaunay triangulation contains several other interesting graphs as subgraphs, such as a minimum spanning tree of a given point set (Exercise 6). In the plane, this leads to an $O(n \log n)$ algorithm for the minimum spanning tree. In \mathbf{R}^3 , subcomplexes of the Delaunay triangulation, the so-called α -complexes, have been successfully used in molecular modeling (see, e.g., Edelsbrunner [Ede98]); they allow one to quickly answer questions such as, “how many tunnels and voids are there in the given molecule?”

Robot motion planning using Voronoi diagrams (or, more generally, the *retraction approach*, where the whole free space for the robot is replaced by some suitable low-dimensional skeleton) was first considered by Ó’Dúnlaig and Yap [ÓY85]. Algorithmic motion planning is an extensive discipline with innumerable variants of the problem. For a brief introduction from the computational-geometric point of view see, e.g., [dBvKOS97]; among several monographs we mention Laumond and Overmars [LO96] and Latombe [Lat91].

The spatial interpolation of functions using Voronoi diagrams was considered by Sibson [Sib81].

Exercises

1. Prove that the region $reg(p)$ of a point p in the Voronoi diagram of a finite point set $P \subset \mathbf{R}^d$ is unbounded if and only if p lies on the surface of $\text{conv}(P)$. \square
2. (a) Show that the Voronoi diagram of the $2n$ -point set $\{(\frac{i}{n}, 0, 0) : i = 1, 2, \dots, n\} \cup \{(0, 1, \frac{j}{n}) : j = 1, 2, \dots, n\}$ in \mathbf{R}^3 has $\Omega(n^2)$ vertices. \square
 (b) Let $d = 2k+1$ be odd, let e_1, \dots, e_d be vectors of the standard orthonormal basis in \mathbf{R}^d , and let e_0 stand for the zero vector. For $i = 0, 1, \dots, k$ and $j = 1, 2, \dots, n$, let $p_{i,j} = e_{2i} + \frac{j}{n}e_{2i+1}$. Prove that for every choice of $j_0, j_1, \dots, j_k \in \{1, 2, \dots, n\}$, there is a point in \mathbf{R}^d for which the nearest points among the $p_{i,j}$ are exactly $p_{0,j_0}, p_{1,j_1}, \dots, p_{k,j_k}$. Conclude that the Voronoi diagram of the $p_{i,j}$ has combinatorial complexity $\Omega(n^k) = \Omega(n^{\lceil d/2 \rceil})$. \square

3. (Voronoi diagram of flats) Let $\varepsilon_1, \dots, \varepsilon_{d-1}$ be small distinct positive numbers and for $i = 1, 2, \dots, d-1$ and $j = 1, 2, \dots, n$, let $F_{i,j}$ be the $(d-2)$ -flat $\{x \in \mathbf{R}^d: x_i = j, x_d = \varepsilon_i\}$. For every choice of $j_1, j_2, \dots, j_{d-1} \in \{1, 2, \dots, n\}$, find a point in \mathbf{R}^d for which the nearest sites (under the Euclidean distance) among the $F_{i,j}$ are exactly $F_{1,j_1}, F_{2,j_2}, \dots, F_{d-1,j_{d-1}}$. Conclude that the Voronoi diagram of the $F_{i,j}$ has combinatorial complexity $\Omega(n^{d-1})$. □
 This example is from Aronov [Aro02].
4. For a finite point set in the plane, define the farthest-point Voronoi diagram as indicated in the text, verify the claimed correspondence with a convex polyhedron in \mathbf{R}^3 , and prove that all nonempty regions are unbounded. □
5. (Delaunay triangulation) Let P be a finite point set in the plane with no 3 points collinear and no 4 points cocircular.
 - (a) Prove that the dual graph of the Voronoi diagram of P , where two points $p, q \in P$ are connected by a straight edge if and only if the boundaries of $\text{reg}(p)$ and $\text{reg}(q)$ share a segment, is a plane graph where the outer face is the complement of $\text{conv}(P)$ and every inner face is a triangle. □
 - (b) Define a graph on P as follows: Two points p and q are connected by an edge if and only if there exists a circular disk with both p and q on the boundary and with no point of P in its interior. Prove that this graph is the same as in (a), and so we have an alternative definition of the Delaunay triangulation. □
6. (Delaunay triangulation and minimum spanning tree) Let $P \subset \mathbf{R}^2$ be a finite point set with no 3 points collinear and no 4 cocircular. Let T be a spanning tree of minimum total edge length in the complete graph with the vertex set P , where the length of an edge is just its Euclidean length. Prove that all edges of T are also edges of the Delaunay triangulation of P . □
7. (Delaunay triangulation by local flipping) Let $P \subset \mathbf{R}^2$ be an n -point set with no 3 points collinear and no 4 cocircular. Let \mathcal{T} be an arbitrary triangulation of $\text{conv}(P)$. Suppose that triangulations $\mathcal{T}_1, \mathcal{T}_2, \dots$ are obtained from \mathcal{T} by successive local flips as described in the notes above (in each step, we select a convex quadrilateral in the current triangulation partitioned into two triangles in a way that is not the Delaunay triangulation of the four vertices and we flip the diagonal of the quadrilateral).
 - (a) Prove that the sequence of triangulations is always finite (and give as good an estimate for its maximum length as you can). □
 - (b) Show that if no local flipping is possible, then the current triangulation is the Delaunay triangulation of P . □
8. Consider a finite set of disjoint segments in the plane. What types of curves may bound the regions in their Voronoi diagram? The region of a

given segment is the set of points for which this segment is a closest one.

\square

9. Let A and B be two finite point sets in the plane. Choose $a_0 \in A$ arbitrarily. Having defined a_0, \dots, a_i and b_1, \dots, b_{i-1} , define b_{i+1} as a point of $B \setminus \{b_1, \dots, b_i\}$ nearest to a_i , and a_{i+1} as a point of $A \setminus \{a_0, \dots, a_i\}$ nearest to b_{i+1} . Continue until one of the sets becomes empty. Prove that at least one of the pairs $(a_i, b_{i+1}), (b_{i+1}, a_{i+1}), i = 0, 1, 2, \dots$, realizes the shortest distance between a point of A and a point of B . (This was used by Eppstein [Epp95] in some dynamical geometric algorithms.) \square
10. (a) Let C be any circle in the plane $x_3 = 0$ (in \mathbf{R}^3). Show that there exists a half-space h such that C is the vertical projection of the set $h \cap U$ onto $x_3 = 0$, where $U = \{x \in \mathbf{R}^3: x_3 = x_1^2 + x_2^2\}$ is the unit paraboloid. \square
 (b) Consider n arbitrary circular disks K_1, \dots, K_n in the plane. Show that there exist only $O(n)$ intersections of their boundaries that lie inside no other K_i (this means that the boundary of the union of the K_i consists of $O(n)$ circular arcs). \square
11. Define a “spherical polytope” as an intersection of n balls in \mathbf{R}^3 (such an object has facets, edges, and vertices similar to an ordinary convex polytope).
 (a) Show that any such spherical polytope in \mathbf{R}^3 has $O(n^2)$ faces. You may assume that the balls are in general position. \square
 (b) Find an example of an intersection of n balls having quadratically many vertices. \square
 (c) Show that the intersection of n *unit* balls has $O(n)$ complexity only. \square

6

Number of Faces in Arrangements

Arrangements of lines in the plane and their higher-dimensional generalization, arrangements of hyperplanes in \mathbf{R}^d , are a basic geometric structure whose significance is comparable to that of convex polytopes. In fact, arrangements and convex polytopes are quite closely related: A cell in a hyperplane arrangement is a convex polyhedron, and conversely, each hyperplane arrangement in \mathbf{R}^d corresponds canonically to a convex polytope in \mathbf{R}^{d+1} of a special type, the so-called zonotope. But as is often the case with different representations of the same mathematical structure, convex polytopes and arrangements of hyperplanes emphasize different aspects of the structure and lead to different questions.

Whenever we have a problem involving a finite point set in \mathbf{R}^d and partitions of the set by hyperplanes, we can use geometric duality, and we obtain a problem concerning a hyperplane arrangement. Arrangements appear in many other contexts as well; for example, some models of molecules give rise to arrangements of spheres in \mathbf{R}^3 , and automatic planning of the motion of a robot among obstacles involves, implicitly or explicitly, arrangements of surfaces in higher-dimensional spaces.

Arrangements of hyperplanes have been investigated for a long time from various points of view. In several classical areas of mathematics one is mainly interested in topological and algebraic properties of the whole arrangement. Hyperplane arrangements are related to such marvelous objects as Lie algebras, root systems, and Coxeter groups. In the theory of *oriented matroids* one studies the systems of sign vectors associated to hyperplane arrangements in an abstract axiomatic setting.

We are going to concentrate on estimating the combinatorial complexity (number of faces) in arrangements and neglect all the other directions.

6.1 Arrangements of Hyperplanes

We recall from Section 4.1 that for a finite set H of lines in the plane, the arrangement of H is a partition of the plane into relatively open convex subsets, the *faces* of the arrangement. In this particular case, the faces are the vertices (0-faces), the edges (1-faces), and the cells (2-faces).¹

An arrangement of a finite set H of hyperplanes in \mathbf{R}^d is again a partition of \mathbf{R}^d into relatively open convex faces. Their dimensions are 0 through d . As in the plane, the 0-faces are called vertices, the 1-faces edges, and the d -faces *cells*. Sometimes the $(d-1)$ -faces are referred to as *facets*.

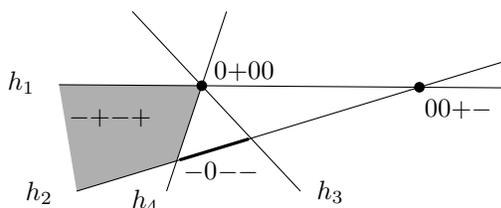
The cells are the connected components of $\mathbf{R}^d \setminus \bigcup H$. To obtain the facets, we consider the $(d-1)$ -dimensional arrangements induced in the hyperplanes of H by their intersections with the other hyperplanes. That is, for each $h \in H$ we take the connected components of $h \setminus \bigcup_{h' \in H: h' \neq h} h'$. To obtain k -faces, we consider every possible k -flat L defined as the intersection of some $d-k$ hyperplanes of H . The k -faces of the arrangement lying within L are the connected components of $L \setminus \bigcup (H \setminus H_L)$, where $H_L = \{h \in H: L \subseteq h\}$.

Remark on sign vectors. A face of the arrangement of H can be described by its *sign vector*. First we need to fix the *orientation* of each hyperplane $h \in H$. Each $h \in H$ partitions \mathbf{R}^d into three regions: h itself and the two open half-spaces determined by it. We choose one of these open half-spaces as positive and denote it by h^\oplus , and we let the other one be negative, denoted by h^\ominus .

Let F be a face of the arrangement of H . We define the *sign vector* of F (with respect to the chosen orientations of the hyperplanes) as $\sigma(F) = (\sigma_h: h \in H)$, where

$$\sigma_h = \begin{cases} +1 & \text{if } F \subseteq h^\oplus, \\ 0 & \text{if } F \subseteq h, \\ -1 & \text{if } F \subseteq h^\ominus. \end{cases}$$

The sign vector determines the face F , since we have $F = \bigcap_{h \in H} h^{\sigma_h}$, where $h^0 = h$, $h^1 = h^\oplus$, and $h^{-1} = h^\ominus$. The following drawing shows the sign vectors of the marked faces in a line arrangement. Only the signs are shown, and the positive half-planes lie above their lines.



¹ This terminology is not unified in the literature. What we call faces are sometimes referred to as cells (0-cells, 1-cells, and 2-cells).

Of course, not all possible sign vectors correspond to nonempty faces. For n lines, there are 3^n sign vectors but only $O(n^2)$ faces, as we will derive below.

Counting the cells in a hyperplane arrangement. We want to count the maximum number of faces in an arrangement of n hyperplanes in \mathbf{R}^d . As we will see, this is much simpler than the similar task for convex polytopes!

If a set H of hyperplanes is in general position, which means that the intersection of every k hyperplanes is $(d-k)$ -dimensional, $k = 2, 3, \dots, d+1$, the arrangement of H is called *simple*. For $|H| \geq d+1$ it suffices to require that every d hyperplanes intersect at a single point and no $d+1$ have a common point.

Every d -tuple of hyperplanes in a simple arrangement determines exactly one vertex, and so a simple arrangement of n hyperplanes has exactly $\binom{n}{d}$ vertices. We now calculate the number of cells; it turns out that the order of magnitude is also n^d for d fixed.

6.1.1 Proposition. *The number of cells (d -faces) in a simple arrangement of n hyperplanes in \mathbf{R}^d equals*

$$\Phi_d(n) = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}. \quad (6.1)$$

First proof. We proceed by induction on the dimension d and the number of hyperplanes n . For $d = 1$ we have a line and n points in it. These divide the line into $n+1$ one-dimensional pieces, and formula (6.1) holds. (The formula is also correct for $n = 0$ and all $d \geq 1$, since the whole space, with no hyperplanes, is a single cell.)

Now suppose that we are in dimension d , we have $n-1$ hyperplanes, and we insert another one. Since we assume general position, the $n-1$ previous hyperplanes divide the newly inserted hyperplane h into $\Phi_{d-1}(n-1)$ cells by the inductive hypothesis. Each such $(d-1)$ -dimensional cell within h partitions one d -dimensional cell into exactly two new cells. The total increase in the number of cells caused by inserting h is thus $\Phi_{d-1}(n-1)$, and so

$$\Phi_d(n) = \Phi_d(n-1) + \Phi_{d-1}(n-1).$$

Together with the initial conditions (for $d = 1$ and for $n = 0$), this recurrence determines all values of Φ , and so it remains to check that formula (6.1) satisfies the recurrence. We have

$$\begin{aligned} \Phi_d(n-1) + \Phi_{d-1}(n-1) &= \binom{n-1}{0} + \left[\binom{n-1}{1} + \binom{n-1}{0} \right] \\ &\quad + \left[\binom{n-1}{2} + \binom{n-1}{1} \right] + \cdots + \left[\binom{n-1}{d} + \binom{n-1}{d-1} \right] \\ &= \binom{n-1}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{d} = \Phi_d(n). \end{aligned}$$

□

Second proof. This proof looks simpler, but a complete rigorous presentation is perhaps somewhat more demanding.

We proceed by induction on d , the case $d = 0$ being trivial. Let H be a set of n hyperplanes in \mathbf{R}^d in general position; in particular, we assume that no hyperplane of H is horizontal and no two vertices of the arrangement have the same vertical level (x_d -coordinate).

Let g be an auxiliary horizontal hyperplane lying below all the vertices. A cell of the arrangement of H either is bounded from below, and in this case it has a unique lowest vertex, or is not bounded from below, and then it intersects g . The number of cells of the former type is the same as the number of vertices, which is $\binom{n}{d}$. The cells of the latter type correspond to the cells in the $(d-1)$ -dimensional arrangement induced within g by the hyperplanes of H , and their number is thus $\Phi_{d-1}(n)$. \square

What is the number of faces of the intermediate dimensions $1, 2, \dots, d-1$ in a simple arrangement of n hyperplanes? This is not difficult to calculate using Proposition 6.1.1 (Exercise 1); the main conclusion is that the total number of faces is $O(n^d)$ for a fixed d .

What about nonsimple arrangements? It turns out that a simple arrangement of n hyperplanes maximizes the number of faces of each dimension among arrangements of n hyperplanes. This can be verified by a perturbation argument, which is considerably simpler than the one for convex polytopes (Lemma 5.5.4), and which we omit.

Bibliography and remarks. The paper of Steiner [Ste26] from 1826 gives formulas for the number of faces in arrangements of lines, circles, planes, and spheres. Of course, his results have been extended in many ways since then (see, e.g., Zaslavsky [Zas75]). An early monograph on arrangements is Grünbaum [Grü72].

The questions considered in the subsequent sections, such as the combinatorial complexity of certain parts of arrangements, have been studied mainly in the last twenty years or so. A recent survey discussing a large part of the material of this chapter and providing many more facts and references is Agarwal and Sharir [AS00].

The algebraic and topological investigation of hyperplane arrangements (both in real and complex spaces) is reflected in the book Orlik and Terao [OT91]. Let us remark that in these areas, one usually considers *central arrangements* of hyperplanes, where all the hyperplanes pass through the origin (and so they are linear subspaces of the underlying vector space). If such a central arrangement in \mathbf{R}^d is intersected with a generic hyperplane not passing through the origin, one obtains a $(d-1)$ -dimensional “affine” arrangement such as those considered by us. The correspondence is bijective, and so these two views of arrangements are not very different, but for many results, the formulation with central arrangements is more elegant.

The correspondence of arrangements to zonotopes is thoroughly explained in Ziegler [Zie94].

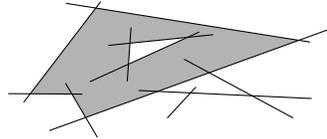
Exercises

1. (a) Count the number of faces of dimensions 1 and 2 for a simple arrangement of n planes in \mathbf{R}^3 . [2]
 (b) Express the number of k -faces in a simple arrangement of n hyperplanes in \mathbf{R}^d . [2]
2. Prove that the number of *unbounded* cells in an arrangement of n hyperplanes in \mathbf{R}^d is $O(n^{d-1})$ (for a fixed d). [2]
3. (a) Check that an arrangement of d or fewer hyperplanes in \mathbf{R}^d has no bounded cell. [2]
 (b) Prove that an arrangement of $d+1$ hyperplanes in general position in \mathbf{R}^d has exactly one bounded cell. [3]
4. How many d -dimensional cells are there in the arrangement of the $\binom{d}{2}$ hyperplanes in \mathbf{R}^d with equations $\{x_i = x_j\}$, where $1 \leq i < j \leq d$? [3]
5. How many d -dimensional cells are there in the arrangement of the hyperplanes in \mathbf{R}^d with the equations $\{x_i - x_j = 0\}$, $\{x_i - x_j = 1\}$, and $\{x_i - x_j = -1\}$, where $1 \leq i < j \leq d$? [5]
6. (Flags in arrangements)
 (a) Let H be a set of n lines in the plane, and let V be the set of vertices of their arrangement. Prove that the number of pairs (v, h) with $v \in V$, $h \in H$, and $v \in h$, i.e., the number of incidences $I(V, H)$, is bounded by $O(n^2)$. (Note that this is trivially true for simple arrangements.) [2]
 (b) Prove that the maximum number of d -tuples (F_0, F_1, \dots, F_d) in an arrangement of n hyperplanes in \mathbf{R}^d , where F_i is an i -dimensional face and F_{i-1} is contained in the closure of F_i , is $O(n^d)$ (d fixed). Such d -tuples are sometimes called *flags* of the arrangement. [3]
7. Let $P = \{p_1, \dots, p_n\}$ be a point set in the plane. Let us say that points x, y have the *same view* of P if the points of P are visible in the same cyclic order from them. If rotating light rays emanate from x and from y , the points of P are lit in the same order by these rays. We assume that neither x nor y is in P and that neither of them can see two points of P in occlusion.
 (a) Show that the maximum possible number of points with mutually distinct views of P is $O(n^4)$. [2]
 (b) Show that the bound in (a) cannot be improved in general. [4]

6.2 Arrangements of Other Geometric Objects

Arrangements can be defined not only for hyperplanes but also for other geometric objects. For example, what is the arrangement of a finite set H of segments in the plane? As in the case of lines, it is a decomposition of the

plane into faces of dimension 0, 1, 2: the vertices, the edges, and the cells, respectively. The vertices are the intersections of the segments, the edges are the portions of the segments after removing the vertices, and the cells (2-faces) are the connected components of $\mathbf{R}^2 \setminus \bigcup H$. (Note that the endpoints of the segments are *not* included among the vertices.) While the cells of line arrangements are convex polygons, those in arrangements of segments can be complicated regions, even with holes:



It is almost obvious that the total number of faces of the arrangement of n segments is at most $O(n^2)$. What is the maximum number of edges on the boundary of a single cell in such an arrangement? This seemingly innocuous question is surprisingly difficult, and most of Chapter ?? revolves around it.

Let us now present the definition of the arrangement for arbitrary sets $A_1, A_2, \dots, A_n \subseteq \mathbf{R}^d$. The arrangement is a subdivision of space into connected pieces again called the *faces*. Each face is an inclusion-maximal connected set that “crosses no boundary.” More precisely, first we define an equivalence relation \approx on \mathbf{R}^d : We put $x \approx y$ whenever x and y lie in the same subcollection of the A_i , that is, whenever $\{i: x \in A_i\} = \{i: y \in A_i\}$. So for each $I \subseteq \{1, 2, \dots, n\}$, we have one possible equivalence class, namely $\{x \in \mathbf{R}^d: x \in A_i \Leftrightarrow i \in I\}$ (this is like a field in the Venn diagram of the A_i). But in typical geometric situations, most of the classes are empty. The faces of the arrangement of the A_i are the connected components of the equivalence classes. The reader is invited to check that for both hyperplane arrangements and arrangements of segments this definition coincides with the earlier ones.

Arrangements of algebraic surfaces. Quite often one needs to consider arrangements of the zero sets of polynomials. Let $p_1(x_1, x_2, \dots, x_d), \dots, p_n(x_1, x_2, \dots, x_d)$ be polynomials with real coefficients in d variables, and let $Z_i = \{x \in \mathbf{R}^d: p_i(x) = 0\}$ be the zero set of p_i . Let D denote the maximum of the degrees of the p_i ; when speaking of the arrangement of Z_1, \dots, Z_n , one usually assumes that D is bounded by some (small) constant. Without a bound on D , even a single Z_i can have arbitrarily many connected components.

In many cases, the Z_i are algebraic surfaces, such as ellipsoids, paraboloids, etc., but since we are in the real domain, sometimes they need not look like surfaces at all. For example, the zero set of the polynomial $p(x_1, x_2) = x_1^2 + x_2^2$ consists of the single point $(0, 0)$. Although it is sometimes convenient to think of the Z_i as surfaces, the results stated below apply to zero sets of arbitrary polynomials of bounded degree.

It is known that if both d and D are considered as constants, the maximum number of faces in the arrangement of Z_1, Z_2, \dots, Z_n as above is at most

$O(n^d)$. This is one of the most useful results about arrangements, with many surprising applications (a few are outlined below and in the exercises). In the literature one often finds a (formally weaker) version dealing with *sign patterns* of the polynomials p_i . A vector $\sigma \in \{-1, 0, +1\}^n$ is called a sign pattern of p_1, p_2, \dots, p_n if there exists an $x \in \mathbf{R}^d$ such that the sign of $p_i(x)$ is σ_i , for all $i = 1, 2, \dots, n$. Trivially, the number of sign patterns for any n polynomials is at most 3^n . For $d = 1$, it is easy to see that the actual number of sign patterns is much smaller, namely at most $2nD + 1$ (Exercise 1). It is not so easy to prove, but still true, that there are at most $C(d, D) \cdot n^d$ sign patterns in dimension d . This kind of result is generally called the *Milnor–Thom theorem* (with some more justice, it should also bear the names of Oleinik and Petrovskii, who proved slightly weaker results much earlier). Here is a more precise (and more recent) version of this statement, where the dependence on D and d is specified quite precisely.

6.2.1 Theorem (Number of sign patterns). *Let p_1, p_2, \dots, p_n be d -variate real polynomials of degree at most D . The number of faces in the arrangement of their zero sets $Z_1, Z_2, \dots, Z_n \subseteq \mathbf{R}^d$, and consequently the number of sign patterns of p_1, \dots, p_n as well is at most $2(2D)^d \sum_{i=0}^d 2^i \binom{4n+1}{i}$. For $n \geq d \geq 2$, this expression is bounded by*

$$\left(\frac{50Dn}{d}\right)^d.$$

Proofs of these results are not included here because they would require at least one more chapter. They belong to the field of *real algebraic geometry*. The classical, deep, and extremely extensive field of *algebraic geometry* mostly studies algebraic varieties over algebraically closed fields, such as the complex numbers (and the questions of combinatorial complexity in our sense are not among its main interests). Real algebraic geometry investigates algebraic varieties and related concepts over the real numbers or other real-closed fields; the presence of ordering and the missing roots of polynomials makes its flavor distinctly different.

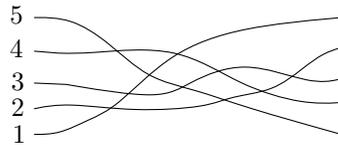
Arrangements of pseudolines. An arrangement of pseudolines is a natural generalization of an arrangement of lines. Lines are replaced by curves, but we insist that these curves behave, in a suitable sense, like lines: For example, no two of them intersect more than once. This kind of generalization is quite different from, say, arrangements of planar algebraic curves, and so it perhaps does not quite belong to the present section. But besides mentioning pseudoline arrangements as a useful and interesting concept, we also need them for a (typical) example of application of Theorem 6.2.1, and so we kill two birds with one stone by discussing them here.

An (*affine*) *arrangement of pseudolines* can be defined as the arrangement of a finite collection of curves in the plane that satisfy the following conditions:

- (i) Each curve is x -monotone and unbounded in both directions; in other words, it intersects each vertical line in exactly one point.
- (ii) Every two of the curves intersect in exactly one point and they cross at the intersection. (We do not permit “parallel” pseudolines, for they would complicate the definition unnecessarily.)²

The curves are called *pseudolines*, but while “being a line” is an absolute notion, “being a pseudoline” makes sense only with respect to a given collection of curves.

Here is an example of a (simple) arrangement of 5 pseudolines:

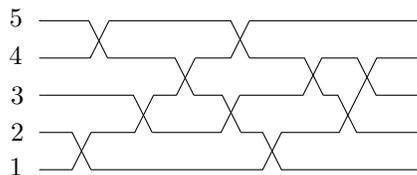


Much of what we have proved for arrangements of lines is true for arrangements of pseudolines as well. This holds for the maximum number of vertices, edges, and cells, but also for more sophisticated results like the Szemerédi–Trotter theorem on the maximum number of incidences of m points and n lines; these results have proofs that do not use any properties of straight lines not shared by pseudolines.

One might be tempted to say that pseudolines are curves that behave topologically like lines, but as we will see below, in at least one sense this is profoundly wrong. The correct statement is that every two of them behave topologically like two lines, but arrangements of pseudolines are more general than arrangements of lines.

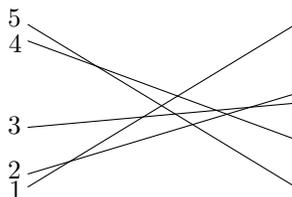
We should first point out that there is no problem with the “local” structure of the pseudolines, since each pseudoline arrangement can be redrawn equivalently (in a sense defined precisely below) by polygonal lines, as a *wiring diagram*:

² This “affine” definition is a little artificial, and we use it only because we do not want to assume the reader’s familiarity with the topology of the projective plane. In the literature one usually considers arrangements of pseudolines in the projective plane, where the definition is very natural: Each pseudoline is a closed curve whose removal does not disconnect the projective plane, and every two pseudolines intersect exactly once (which already implies that they cross at the intersection point). Moreover, one often adds the condition that the curves do not form a single *pencil*; i.e., not all of them have a common point, since otherwise, one would have to exclude the case of a pencil in the formulation of many theorems. But here we are not going to study pseudoline arrangements in any depth.



The difference between pseudoline arrangements and line arrangements is of a more global nature.

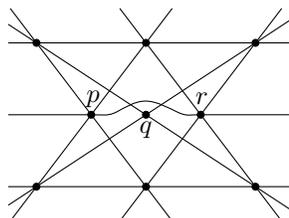
The arrangement of 5 pseudolines drawn above can be realized by straight lines:



What is the meaning of “realization by straight lines”? To this end, we need a suitable notion of equivalence of two arrangements of pseudolines. There are several technically different possibilities; we again use an “affine” notion, one that is very simple to state but not the most common. Let H be a collection of n pseudolines. We number the pseudolines $1, 2, \dots, n$ in the order in which they appear on the left of the arrangement, say from the bottom to the top. For each i , we write down the numbers of the other pseudolines in the order they are encountered along the pseudoline i from left to right. For a simple arrangement we obtain a permutation π_i of $\{1, 2, \dots, n\} \setminus \{i\}$ for each i . For the arrangement in the pictures, we have $\pi_1 = (2, 3, 5, 4)$, $\pi_2 = (1, 5, 4, 3)$, $\pi_3 = (1, 5, 4, 2)$, $\pi_4 = (5, 1, 3, 2)$, and $\pi_5 = (4, 1, 3, 2)$. For a nonsimple arrangement, some of the π_i are linear *quasiorderings*, meaning that several consecutive numbers can be chunked together. We call two arrangements *affinely isomorphic* if they yield the same π_1, \dots, π_n , i.e., if each pseudoline meets the others in the same (quasi)order as the corresponding pseudoline in the other arrangement. Two affinely isomorphic pseudoline arrangements can be converted one to another by a suitable homeomorphism of the plane.³

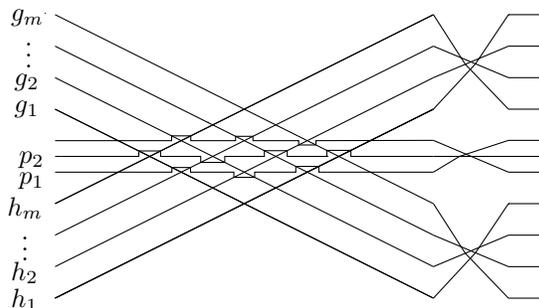
³ The more usual notion of *isomorphism* of pseudoline arrangements is defined for arrangements in the projective plane. The arrangement of H is isomorphic to the arrangement of H' if there exists a homeomorphism φ of the projective plane onto itself such that each pseudoline $h \in H$ is mapped to a pseudoline $\varphi(h) \in H'$. For affinely isomorphic arrangements in the affine plane, the corresponding arrangements in the projective plane are isomorphic, but the isomorphism in the projective plane also allows for mirror reflection and for “relocating the infinity.” Combinatorially, the isomorphism in the projective plane can be described using the (quasi)orderings π_1, \dots, π_n as well. Here the π_i have to agree only up to

An arrangement of pseudolines is *stretchable* if it is affinely isomorphic to an arrangement of straight lines.⁴ It turns out that all arrangements of 8 or fewer pseudolines are stretchable, but there exists a nonstretchable arrangement of 9 pseudolines:



The proof of nonstretchability is based on the *Pappus theorem* in projective geometry, which states that if 8 straight lines intersect as in the drawing, then the points p , q , and r are collinear. By modifying this arrangement suitably, one can obtain a simple nonstretchable arrangement of 9 pseudolines as well.

Next, we show that most of the simple pseudoline arrangements are nonstretchable. The following construction shows that the number of isomorphism classes of simple arrangements of n pseudolines is at least $2^{\Omega(n^2)}$:



We have $m \approx \frac{n}{3}$, and the lines h_1, \dots, h_m and g_1, \dots, g_m form a regular grid. Each of the about $\frac{n}{3}$ pseudolines p_i in the middle passes near $\Omega(n)$ vertices of this grid, and for each such vertex it has a choice of going below it or above. This gives $2^{\Omega(n^2)}$ possibilities in total.

Now we use Theorem 6.2.1 to estimate the number of nonisomorphic simple arrangements of n straight lines. Let the lines be ℓ_1, \dots, ℓ_n , where ℓ_i has the equation $y = a_i x + b_i$ and $a_1 > a_2 > \dots > a_n$. The x -coordinate

a possible reversal and cyclic shift for each i , and also the numbering of the pseudolines by $1, 2, \dots, n$ is not canonical.

We also remark that two arrangements of *lines* are isomorphic if and only if the dual point configurations have the same order type, up to a mirror reflection of the whole configuration (order types are discussed in Section ??).

⁴ For isomorphism in the projective plane, one gets an equivalent notion of stretchability.

of the intersection $\ell_i \cap \ell_j$ is $\frac{b_i - b_j}{a_j - a_i}$. To determine the ordering π_i of the intersections along ℓ_i , it suffices to know the ordering of the x -coordinates of these intersections, and this can be inferred from the signs of the polynomials $p_{ijk}(a_i, b_i, a_j, b_j, a_k, b_k) = (b_i - b_j)(a_k - a_i) - (b_i - b_k)(a_j - a_i)$. So the number of nonisomorphic arrangements of n lines is no larger than the number of possible sign patterns of the $O(n^3)$ polynomials p_{ijk} in the $2n$ variables $a_1, b_1, \dots, a_n, b_n$, and Theorem 6.2.1 yields the upper bound of $2^{O(n \log n)}$. For large n , this is a negligible fraction of the total number of simple pseudoline arrangements. (Similar considerations apply to nonsimple arrangements as well.)

The problem of deciding the stretchability of a given pseudoline arrangement has been shown to be algorithmically difficult (at least NP-hard). One can easily encounter this problem when thinking about line arrangements and drawing pictures: What we draw by hand are really pseudolines, not lines, and even with the help of a ruler it may be almost impossible to decide experimentally whether a given arrangement can really be drawn with straight lines. But there are computational methods that can decide stretchability in reasonable time at least for moderate numbers of lines.

Bibliography and remarks. A comprehensive account of real algebraic geometry is Bochnak, Coste, and Roy [BCR98], and a more recent treatment stressing an algorithmic approach is Basu, Pollack, and Roy [BPR03]. Among the many available introductions to the “classical” algebraic geometry we mention the lively book Cox, Little, and O’Shea [CLO92].

The first bounds in the spirit of Theorem 6.2.1 were obtained by Oleinik and Petrovskii; in particular, Oleinik in her 1951 paper [Ole51] showed that the maximum number of connected components, or more generally the sum of the Betti numbers, of the zero set $V \subset \mathbf{R}^d$ of a polynomial of degree D is at most $D(D-1)^{d-1}$, assuming that V is a smooth hypersurface. Later Milnor [Mil64] proved the slightly worse bound of $D(2D-1)^{d-1}$ with V an arbitrary real algebraic set, that is, the set of common zeros in \mathbf{R}^d of polynomials of degree at most D . Similar results were obtained independently by Thom [Tho65]. Milnor also established a bound for a basic semialgebraic set, that is, a set defined by a conjunction of polynomial inequalities. Bounds on the number of sign patterns and cells in arrangements, which were considered only later, can be derived from the Milnor–Thom results by elementary reductions (see, e.g., the seminal paper of Ben-Or [BO83], where lower bounds on the complexity of a number of basic algorithmic problems were derived, or Alon [Alo86]). Warren [War68] showed that the number of d -dimensional cells in the arrangement as in Theorem 6.2.1 is at most $2(2D)^d \sum_{i=0}^d 2^i \binom{n}{i}$. The extension of Warren’s bounds to faces of all dimensions was obtained by Pollack and Roy [PR93].

Sometimes we have polynomials in many variables, but we are interested only in sign patterns attained at points that satisfy some additional algebraic conditions. Such a situation is covered by a result of Basu, Pollack, and Roy [BPR96]: The number of sign patterns attained by n polynomials of degree at most D on a k -dimensional algebraic variety $V \subseteq \mathbf{R}^d$, where V can be defined by polynomials of degree at most D , is at most $\binom{n}{k} O(D)^d$.

While bounding the number of sign patterns of multivariate polynomials appears complicated, there is a beautiful short proof of an almost tight bound on the number of *zero patterns*, due to Rónyai, Babai, and Ganapathy [RBG01], which we now sketch (in the simplest form, giving a slightly suboptimal result). A vector $\zeta \in \{0, 1\}^n$ is a zero pattern of d -variate polynomials p_1, \dots, p_n with coefficients in a field F if there exists an $x = x(\zeta) \in F^d$ with $p_i(x) = 0$ exactly for the i with $\zeta_i = 0$. We show that if all the p_i have degree at most D , then the number of zero patterns cannot exceed $\binom{Dn+d}{d}$. For each zero pattern ζ , let q_ζ be the polynomial $\prod_{i: \zeta_i \neq 0} p_i$. We have $\deg q_\zeta \leq Dn$. Let us consider the q_ζ as elements of the vector space L of all d -variate polynomials over F of degree at most Dn . Using the basis of L consisting of all monomials of degree at most Dn , we obtain $\dim L \leq \binom{Dn+d}{d}$. It remains to verify that the q_ζ are linearly independent (assuming that no p_i is identically 0). Suppose that $\sum_\zeta \alpha_\zeta q_\zeta = 0$ with $\alpha_\zeta \in F$ not all 0. Choose a zero pattern ξ with $\alpha_\xi \neq 0$ and with the largest possible number of 0's, and substitute $x(\xi)$ into $\sum_\zeta \alpha_\zeta q_\zeta$. This yields $\alpha_\xi = 0$, a contradiction.

Pseudoline arrangements. The founding paper is Levi [Lev26], where, among others, the nonstretchable arrangement of 9 lines drawn above was presented. A concise survey was written by Goodman [Goo97].

Pseudoline arrangements, besides being very natural, have also turned out to be a fruitful generalization of line arrangements. Some problems concerning line arrangements or point configurations were first solved only in the more general setting of pseudoline arrangements, and certain algorithms for line arrangements, the so-called *topological sweep* methods, use an auxiliary pseudoline to speed up the computation; see [Goo97].

Infinite families of pseudolines have been considered as well, and even *topological planes*, which are analogues of the projective plane but made of pseudolines. It is known that every finite configuration of pseudolines can be extended to a topological plane, and there are uncountably many distinct topological planes; see Goodman, Pollack, Wenger, and Zamfirescu [GPWZ94].

Oriented matroids. The possibility of representing each pseudoline arrangement by a wiring diagram makes it clear that a pseudoline arrangement can also be considered as a purely combinatorial object.

The appropriate combinatorial counterpart of a pseudoline arrangement is called an oriented matroid of rank 3. More generally, similar to arrangements of pseudolines, one can define arrangements of *pseudohyperplanes* in \mathbf{R}^d , and these are combinatorially captured by oriented matroids of rank $d+1$. Here the rank is one higher than the space dimension, because an oriented matroid of rank d is usually viewed as a combinatorial abstraction of a *central* arrangement of hyperplanes in \mathbf{R}^d (with all hyperplanes passing through 0).

There are several different but equivalent definitions of an oriented matroid. We present a definition in the so-called *covector form*. An *oriented matroid* is a set $\mathcal{V}^* \subseteq \{-1, 0, 1\}^n$ that is symmetric ($v \in \mathcal{V}^*$ implies $-v \in \mathcal{V}^*$), contains the zero vector, and satisfies the following two more complicated conditions:

- (Closed under composition) If $u, v \in \mathcal{V}^*$, then $u \circ v \in \mathcal{V}^*$, where $(u \circ v)_i = u_i$ if $u_i \neq 0$ and $(u \circ v)_i = v_i$ if $u_i = 0$.
- (Admits elimination) If $u, v \in \mathcal{V}^*$ and $j \in S(u, v) = \{i: u_i = -v_i \neq 0\}$, then there exists $w \in \mathcal{V}^*$ such that $w_j = 0$ and $w_i = (u \circ v)_i$ for all $i \notin S(u, v)$.

The *rank* of an oriented matroid \mathcal{V}^* is the largest r such that there is an increasing chain $v_1 \prec v_2 \prec \cdots \prec v_r$, $v_i \in \mathcal{V}^*$, where $u \preceq v$ means $u_i \preceq v_i$ for all i and where $0 \prec 1$ and $0 \prec -1$. At first sight, all this may look quite mysterious, but it becomes much clearer if one thinks of a basic example, where \mathcal{V}^* is the set of sign vectors of all faces of a central arrangement of hyperplanes in \mathbf{R}^d .

It turns out that every oriented matroid of rank 3 corresponds to an arrangement of pseudolines. More generally, *Lawrence's representation theorem* asserts that every oriented matroid of rank d comes from some central arrangement of pseudohyperplanes in \mathbf{R}^d , and so the purely combinatorial notion of oriented matroid corresponds, essentially uniquely, to the topological notion of a (central) arrangement of pseudohyperplanes.⁵

Oriented matroids are also naturally obtained from configurations of points or vectors. In the notation of Section ?? (Gale transform), if $\bar{\mathbf{a}}$ is a sequence of n vectors in \mathbf{R}^r , then both the sets $\text{sgn}(\text{LinVal}(\bar{\mathbf{a}}))$ and $\text{sgn}(\text{LinDep}(\bar{\mathbf{a}}))$ are oriented matroids in the sense of the above definition. The first one has rank r , and the second, rank $n-r$.

We are not going to say much more about oriented matroids, referring to Ziegler [Zie94] for a quick introduction and to Björner, Las Vergnas, Sturmfels, White, and Ziegler [BVS⁺99] for a comprehensive account.

⁵ The correspondence need not really be one-to-one. For example, the oriented matroids of two projectively isomorphic pseudoline arrangements agree only up to reorientation.

Stretchability. The following results illustrate the surprising difficulty of the stretchability problem for pseudoline arrangements. They are analogous to the statements about realizability of 4-dimensional convex polytopes mentioned in Section 5.3, and they were actually found much earlier.

Certain (simple) stretchable arrangements of n pseudolines require coefficients with $2^{\Omega(n)}$ digits in the equations of the lines, in every straight-line realization (Goodman, Pollack, and Sturmfels [GPS90]). Deciding the stretchability of a given pseudoline arrangement is NP-hard (Shor [Sho91] has a relatively simple proof), and in fact, it is polynomially equivalent to the problem of solvability of a system of polynomial inequalities with integer coefficients. This follows from results of Mnëv, published in Russian in 1985 (proofs were only sketched; see [Mne89] for an English version). This work went unnoticed in the West for some time, and so some of the results were rediscovered by other authors.

Although detailed proofs of such theorems are technically demanding, the principle is rather simple. Given two real numbers, suitably represented by geometric quantities, one can produce their sum and their product by classical geometric constructions by ruler. (Since ruler constructions are invariant under projective transformations, the numbers are represented as cross-ratios.) By composing such constructions, one can express the solvability of $p(x_1, \dots, x_n) = 0$, for a given n -variate polynomial p with integer coefficients, by the stretchability of a suitable arrangement in the projective plane. Dealing with inequalities and passing to simple arrangements is somewhat more complicated, but the idea is similar.

Practical algorithms for deciding stretchability have been studied extensively by Bokowski and Sturmfels [BS89] and by Richter-Gebert (see, e.g., [RG99]).

Mnëv [Mne89] was mainly interested in the *realization spaces* of arrangements. Let H be a fixed stretchable arrangement. Each straight-line arrangement H' affinely isomorphic to H can be represented by a point in \mathbf{R}^{2n} , with the $2n$ coordinates specifying the coefficients in the equations of the lines of H' . Considering all possible H' for a given H , we obtain a subset of \mathbf{R}^{2n} . For some time it was conjectured that this set, the realization space of H , has to be path-connected, which would mean that one straight-line realization could be converted to any other by a continuous motion while retaining the affine isomorphism type.⁶ Not only is this false, but the realization space can have arbitrarily many components. In a suitable sense, it can even have

⁶ In fact, these questions have been studied mainly for the isomorphism of arrangements in the projective plane. There one has to be a little careful, since a mirror reflection can easily make the realization space disconnected, and so the mirror reflection (or the whole action of the general linear group) is factored out first.

arbitrary topological type. Whenever $A \subseteq \mathbf{R}^n$ is a set definable by a formula involving finitely many polynomial inequalities with integer coefficients, Boolean connectives, and quantifiers, there is a line arrangement whose realization space S is homotopy equivalent to A (Mnëv's main result actually talks about the stronger notion of *stable equivalence* of S and A ; see, e.g., [Goo97] or [BVS⁺99]). Similar theorems were proved by Richter-Gebert for the realization spaces of 4-dimensional polytopes [RG99], [RG97].

These results for arrangements and polytopes can be regarded as instances of a vague but probably quite general principle: “*Almost none of the combinatorially imaginable geometric configurations are geometrically realizable, and it is difficult to decide which ones are.*” Of course, there are exceptions, such as the graphs of 3-dimensional convex polytopes.

Encoding pseudoline arrangements. The lower bound $2^{\Omega(n^2)}$ for the number of isomorphism classes of pseudoline arrangements is asymptotically tight. Felsner [Fel97] found a nice encoding of such an arrangement by an $n \times n$ matrix of 0's and 1's, from which the isomorphism type can be reconstructed: The entry (i, j) of the matrix is 1 iff the j th leftmost crossing along the pseudoline number i is with a pseudoline whose number k is larger than i .

Exercises

1. Let $p_1(x), \dots, p_n(x)$ be univariate real polynomials of degree at most D . Check that the number of sign patterns of the p_i is at most $2nD+1$. □
2. (Intersection graphs) Let S be a set of n line segments in the plane. The *intersection graph* of S is the graph on n vertices, which correspond to the segments of S , with two vertices connected by an edge if and only if the corresponding two segments intersect.
 - (a) Prove that the graph obtained from K_5 by subdividing each edge exactly once is not the intersection graph of segments in the plane (and not even the intersection graph of any arcwise connected sets in the plane). □
 - (b) Use Theorem 6.2.1 to prove that *most* graphs are not intersection graphs of segments: While the total number of graphs on n given vertices is $2^{\binom{n}{2}} = 2^{n^2/2+O(n)}$, only $2^{O(n \log n)}$ of them are intersection graphs of segments (be careful about collinear segments!). □
 - (c) Show that the number of (isomorphism classes of) intersection graphs of planar arcwise connected sets, and even of planar convex sets, on n vertices is much larger: $2^{\Omega(n^2)}$. (The right order of magnitude in both of these cases is $2^{(3/8+o(1))n^2}$, as was determined by Pach and Tóth [PT03].) □

3. (Number of combinatorially distinct simplicial convex polytopes) Use Theorem 6.2.1 to prove that for every dimension $d \geq 3$ there exists $C_d > 0$ such that the number of combinatorial types of *simplicial* polytopes in \mathbf{R}^d with n vertices is at most $2^{C_d n \log n}$. (The combinatorial equivalence means isomorphic face lattices; see Definition 5.3.3.) \square
- Such a result was proved by Alon [Alo86] and by Goodman and Pollack [GP86].
4. (Sign patterns of matrices and rank) Let A be a real $n \times n$ matrix. The *sign matrix* $\sigma(A)$ is the $n \times n$ matrix with entries in $\{-1, 0, +1\}$ given by the signs of the corresponding entries in A .
- (a) Check that A has rank at most q if and only if there exist $n \times q$ matrices U and V with $A = UV^T$. \square
- (b) Estimate the number of distinct sign matrices of rank q using Theorem 6.2.1, and conclude that there exists an $n \times n$ matrix S containing only entries $+1$ and -1 such that any real matrix A with $\sigma(A) = S$ has rank at least cn , with a suitable constant $c > 0$. \square
- The result in (b) is from Alon, Frankl, and Rödl [AFR85] (for another application see [Mat96]).

Bibliography

The references are sorted alphabetically by the abbreviations (rather than by the authors' names).

- [ABS97] D. Avis, D. Bremner, and R. Seidel. How good are convex hull algorithms? *Comput. Geom. Theory Appl.*, 7:265–302, 1997.
- [AF92] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete Comput. Geom.*, 8:295–313, 1992.
- [AFR85] N. Alon, P. Frankl, and V. Rödl. Geometrical realization of set systems and probabilistic communication complexity. In *Proc. 26th IEEE Symposium on Foundations of Computer Science*, pages 277–280, 1985.
- [AG86] N. Alon and E. Györi. The number of small semispaces of a finite set of points in the plane. *J. Combin. Theory Ser. A*, 41:154–157, 1986.
- [AI88] F. Aurenhammer and H. Imai. Geometric relations among Voronoi diagrams. *Geom. Dedicata*, 27:65–75, 1988.
- [AK85] N. Alon and G. Kalai. A simple proof of the upper bound theorem. *European J. Combin.*, 6:211–214, 1985.
- [AK00] F. Aurenhammer and R. Klein. Voronoi diagrams. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 201–290. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 2000.
- [Alo86] N. Alon. The number of polytopes, configurations, and real matroids. *Mathematika*, 33:62–71, 1986.
- [AMS94] B. Aronov, J. Matoušek, and M. Sharir. On the sum of squares of cell complexities in hyperplane arrangements. *J. Combin. Theory Ser. A*, 65:311–321, 1994.
- [AMS98] P. K. Agarwal, J. Matoušek, and O. Schwarzkopf. Computing many faces in arrangements of lines and segments. *SIAM J. Comput.*, 27(2):491–505, 1998.
- [APS93] B. Aronov, M. Pellegrini, and M. Sharir. On the zone of a surface in a hyperplane arrangement. *Discrete Comput. Geom.*, 9(2):177–186, 1993.

- [Aro02] B. Aronov. A lower bound on Voronoi diagram complexity. *Inf. Process. Lett.*, 83:183–185, 2002.
- [AS00] P. K. Agarwal and M. Sharir. Arrangements and their applications. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 49–119. North-Holland, Amsterdam, 2000.
- [Aur91] F. Aurenhammer. Voronoi diagrams: A survey of a fundamental geometric data structure. *ACM Comput. Surv.*, 23(3):345–405, September 1991.
- [BCM99] H. Brönnimann, B. Chazelle, and J. Matoušek. Product range spaces, sensitive sampling, and derandomization. *SIAM J. Comput.*, 28:1552–1575, 1999.
- [BCR98] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*. Springer, Berlin etc., 1998. Transl. from the French, revised and updated edition.
- [BEPY91] M. Bern, D. Eppstein, P. Plassman, and F. Yao. Horizon theorems for lines and polygons. In J. Goodman, R. Pollack, and W. Steiger, editors, *Discrete and Computational Geometry: Papers from the DIMACS Special Year*, volume 6 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 45–66. American Mathematical Society, Association for Computing Machinery, Providence, RI, 1991.
- [BK63] W. Bonnice and V. L. Klee. The generation of convex hulls. *Math. Ann.*, 152:1–29, 1963.
- [BO83] M. Ben-Or. Lower bounds for algebraic computation trees. In *Proc. 15th Annu. ACM Sympos. Theory Comput.*, pages 80–86, 1983.
- [BPR96] S. Basu, R. Pollack, and M.-F. Roy. On the number of cells defined by a family of polynomials on a variety. *Mathematika*, 43:120–126, 1996.
- [BPR03] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*. Algorithms and Computation in Mathematics 10. Springer, Berlin, 2003.
- [Bre93] G. Bredon. *Topology and Geometry (Graduate Texts in Mathematics 139)*. Springer-Verlag, Berlin etc., 1993.
- [Brø83] A. Brønsted. *An Introduction to Convex Polytopes*. Springer-Verlag, New York, NY, 1983.
- [BS89] J. Bokowski and B. Sturmfels. *Computational Synthetic Geometry*. Lect. Notes in Math. 1355. Springer-Verlag, Heidelberg, 1989.
- [BV82] E. O. Buchman and F. A. Valentine. Any new Helly numbers? *Amer. Math. Mon.*, 89:370–375, 1982.
- [BVS⁺99] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler. *Oriented Matroids (2nd edition)*. Encyclopedia

- of Mathematics 46. Cambridge University Press, Cambridge, 1999.
- [Car07] C. Carathéodory. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. *Math. Ann.*, 64:95–115, 1907.
- [CEG⁺90] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl. Combinatorial complexity bounds for arrangements of curves and spheres. *Discrete Comput. Geom.*, 5:99–160, 1990.
- [CEG⁺93] B. Chazelle, H. Edelsbrunner, L. Guibas, M. Sharir, and J. Snoeyink. Computing a face in an arrangement of line segments and related problems. *SIAM J. Comput.*, 22:1286–1302, 1993.
- [CEGS89] B. Chazelle, H. Edelsbrunner, L. Guibas, and M. Sharir. A singly-exponential stratification scheme for real semi-algebraic varieties and its applications. In *Proc. 16th Internat. Colloq. Automata Lang. Program.*, volume 372 of *Lecture Notes Comput. Sci.*, pages 179–192. Springer-Verlag, Berlin etc., 1989.
- [CEM⁺96] K. L. Clarkson, D. Eppstein, G. L. Miller, C. Sturivant, and S.-H. Teng. Approximating center points with iterative Radon points. *Internat. J. Comput. Geom. Appl.*, 6:357–377, 1996.
- [CF90] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.
- [CGL85] B. Chazelle, L. J. Guibas, and D. T. Lee. The power of geometric duality. *BIT*, 25:76–90, 1985.
- [Cha93a] B. Chazelle. Cutting hyperplanes for divide-and-conquer. *Discrete Comput. Geom.*, 9(2):145–158, 1993.
- [Cha93b] B. Chazelle. An optimal convex hull algorithm in any fixed dimension. *Discrete Comput. Geom.*, 10:377–409, 1993.
- [Cha00a] T. M. Chan. On levels in arrangements of curves. In *Proc. 41st IEEE Symposium on Foundations of Computer Science*, pages 219–227, 2000.
- [Cha00b] T. M. Chan. Random sampling, halfspace range reporting, and construction of ($\leq k$)-levels in three dimensions. *SIAM J. Comput.*, 30(2):561–575, 2000.
- [Cha00c] B. Chazelle. *The Discrepancy Method*. Cambridge University Press, Cambridge, 2000.
- [Cla88a] K. L. Clarkson. Applications of random sampling in computational geometry, II. In *Proc. 4th Annu. ACM Sympos. Comput. Geom.*, pages 1–11, 1988.
- [Cla88b] K. L. Clarkson. A randomized algorithm for closest-point queries. *SIAM J. Comput.*, 17:830–847, 1988.
- [CLO92] D. Cox, J. Little, and D. O’Shea. *Ideals, Varieties, and Algorithms*. Springer-Verlag, New York, NY, 1992.

- [CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.
- [dBvKOS97] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, Berlin, 1997.
- [Del34] B. Delaunay. Sur la sphère vide. A la memoire de Georges Voronoi. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk*, 7:793–800, 1934.
- [DGK63] L. Danzer, B. Grünbaum, and V. Klee. Helly’s theorem and its relatives. In *Convexity*, volume 7 of *Proc. Symp. Pure Math.*, pages 101–180. American Mathematical Society, Providence, 1963.
- [Dir50] G. L. Dirichlet. Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *J. Reine Angew. Math.*, 40:209–227, 1850.
- [Dol92] V. L. Dolnikov. A generalization of the ham sandwich theorem. *Mat. Zametki*, 52(2):27–37, 1992. In Russian; English translation in *Math. Notes* 52,2:771–779, 1992.
- [Eck93] J. Eckhoff. Helly, Radon and Carathéodory type theorems. In P. M. Gruber and J. M. Wills, editors, *Handbook of Convex Geometry*. North-Holland, Amsterdam, 1993.
- [Ede98] H. Edelsbrunner. Geometry of modeling biomolecules. In P. K. Agarwal, L. E. Kavradi, and M. Mason, editors, *Proc. Workshop Algorithmic Found. Robot.* A. K. Peters, Natick, MA, 1998.
- [EKZ03] D. Eppstein, G. Kuperberg, and G. M. Ziegler. Fat 4-polytopes and fatter 3-spheres. In *Discrete Geometry: In honor of W. Kuperberg’s 60th birthday* (A. Bezdek, ed.), Pure and Applied Mathematics Vol. 253, pages 239–265. Marcel Dekker Inc., New York, NY, 2003.
- [EOS86] H. Edelsbrunner, J. O’Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.*, 15:341–363, 1986.
- [Epp95] D. Eppstein. Dynamic Euclidean minimum spanning trees and extrema of binary functions. *Discrete Comput. Geom.*, 13:111–122, 1995.
- [ES96] H. Edelsbrunner and N. R. Shah. Incremental topological flipping works for regular triangulations. *Algorithmica*, 15:223–241, 1996.
- [ESS93] H. Edelsbrunner, R. Seidel, and M. Sharir. On the zone theorem for hyperplane arrangements. *SIAM J. Comput.*, 22(2):418–429, 1993.

-
- [Far94] G. Farkas. Applications of Fourier's mechanical principle (in Hungarian). *Math. Termés. Értésítő*, 12:457–472, 1893/94. German translation in *Math. Nachr. Ungarn* 12:1–27, 1895.
- [Fel97] S. Felsner. On the number of arrangements of pseudolines. *Discrete Comput. Geom.*, 18:257–267, 1997.
- [Gal63] D. Gale. Neighborly and cyclic polytopes. In V. Klee, editor, *Convexity*, volume 7 of *Proc. Symp. Pure Math.*, pages 225–232. American Mathematical Society, 1963.
- [GGL95] R. L. Graham, M. Grötschel, and L. Lovász, editors. *Handbook of Combinatorics*. North-Holland, Amsterdam, 1995.
- [GJ00] E. Gawrilow and M. Joswig. `polymake`: a framework for analyzing convex polytopes. In G. Kalai and G. M. Ziegler, editors, *Polytopes—Combinatorics and Computation*, pages 43–74. Birkhäuser, Basel, 2000. Software available at <http://www.math.tu-berlin.de/diskregeom/polymake/>.
- [GO97] J. E. Goodman and J. O'Rourke, editors. *Handbook of Discrete and Computational Geometry*. CRC Press LLC, Boca Raton, FL, 1997.
- [Goo97] J. E. Goodman. Pseudoline arrangements. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 83–110. CRC Press LLC, Boca Raton, FL, 1997.
- [GP84] J. E. Goodman and R. Pollack. On the number of k -subsets of a set of n points in the plane. *J. Combin. Theory Ser. A*, 36:101–104, 1984.
- [GP86] J. E. Goodman and R. Pollack. Upper bounds for configurations and polytopes in \mathbb{R}^d . *Discrete Comput. Geom.*, 1:219–227, 1986.
- [GPS90] J. E. Goodman, R. Pollack, and B. Sturmfels. The intrinsic spread of a configuration in \mathbb{R}^d . *J. Amer. Math. Soc.*, 3:639–651, 1990.
- [GPWZ94] J. E. Goodman, R. Pollack, R. Wenger, and T. Zamfirescu. Arrangements and topological planes. *Amer. Math. Monthly*, 101(10):866–878, 1994.
- [Grü67] B. Grünbaum. *Convex Polytopes*. John Wiley & Sons, New York, NY, 1967.
- [Grü72] B. Grünbaum. *Arrangements and Spreads*. Regional Conf. Ser. Math. American Mathematical Society, Providence, RI, 1972.
- [Grü03] B. Grünbaum. *Convex Polytopes. Second edition prepared by V. Kaibel, V. Klee, and G. M. Ziegler*. Springer, Berlin, 2003.
- [GW93] P. M. Gruber and J. M. Wills, editors. *Handbook of Convex Geometry (volumes A and B)*. North-Holland, Amsterdam, 1993.
- [JM94] S. Jadhav and A. Mukhopadhyay. Computing a centerpoint of a finite planar set of points in linear time. *Discrete Comput. Geom.*, 12:291–312, 1994.

- [Kat78] M. Katchalski. A Helly type theorem for convex sets. *Can. Math. Bull.*, 21:121–123, 1978.
- [Kir03] P. Kirchberger. Über Tschebyscheffsche Annäherungsmethoden. *Math. Ann.*, 57:509–540, 1903.
- [Kle53] V. Klee. The critical set of a convex body. *Amer. J. Math.*, 75:178–188, 1953.
- [Kle64] V. Klee. On the number of vertices of a convex polytope. *Canadian J. Math.*, 16:701–720, 1964.
- [Kle89] R. Klein. *Concrete and Abstract Voronoi Diagrams*, volume 400 of *Lecture Notes Comput. Sci.* Springer-Verlag, Berlin etc., 1989.
- [Kol01] V. Koltun. Almost tight upper bounds for vertical decompositions in four dimensions. In *Proc. 42nd IEEE Symposium on Foundations of Computer Science*, 2001.
- [KP01] V. Kaibel and M. E. Pfetsch. Computing the face lattice of a polytope from its vertex-facet incidences. Technical Report, Inst. für Mathematik, TU Berlin, 2001.
- [KZ00] G. Kalai and G. M. Ziegler, editors. *Polytopes—Combinatorics and Computation. DMV-seminar Oberwolfach, Germany, November 1997.* Birkhäuser, Basel, 2000.
- [Lat91] J.-C. Latombe. *Robot Motion Planning.* Kluwer Academic Publishers, Boston, 1991.
- [Lee82] D. T. Lee. On k -nearest neighbor Voronoi diagrams in the plane. *IEEE Trans. Comput.*, C-31:478–487, 1982.
- [Lev26] F. Levi. Die Teilung der projektiven Ebene durch Gerade oder Pseudogerade. *Ber. Math.-Phys. Kl. sächs. Akad. Wiss. Leipzig*, 78:256–267, 1926.
- [LMS94] C.-Y. Lo, J. Matoušek, and W. L. Steiger. Algorithms for ham-sandwich cuts. *Discrete Comput. Geom.*, 11:433, 1994.
- [LO96] J.-P. Laumond and M. H. Overmars, editors. *Algorithms for Robotic Motion and Manipulation.* A. K. Peters, Wellesley, MA, 1996.
- [Mat96] J. Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.*, 93:333–344, 1996.
- [McM70] P. McMullen. The maximal number of faces of a convex polytope. *Mathematika*, 17:179–184, 1970.
- [Mil64] J. W. Milnor. On the Betti numbers of real algebraic varieties. *Proc. Amer. Math. Soc.*, 15:275–280, 1964.
- [Mne89] M. E. Mnev. The universality theorems on the classification problem of configuration varieties and convex polytopes varieties. In O. Y. Viro, editor, *Topology and Geometry—Rohlin Seminar*, volume 1346 of *Lecture Notes Math.*, pages 527–544. Springer, Berlin etc., 1989.

-
- [MS71] P. McMullen and G. C. Shephard. *Convex Polytopes and the Upper Bound Conjecture*, volume 3 of *Lecture Notes*. Cambridge University Press, Cambridge, England, 1971.
- [Mul93] K. Mulmuley. *Computational Geometry: An Introduction Through Randomized Algorithms*. Prentice Hall, Englewood Cliffs, NJ, 1993.
- [OBS92] A. Okabe, B. Boots, and K. Sugihara. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley & Sons, Chichester, UK, 1992.
- [Ole51] O. A. Oleinik. Estimates of the Betti numbers of real algebraic hypersurfaces (in Russian). *Mat. Sbornik (N. S.)*, 28(70):635–640, 1951.
- [OT91] P. Orlik and H. Terao. *Arrangements of Hyperplanes*. Springer-Verlag, Berlin etc., 1991.
- [ÓY85] C. Ó’Dúnlaing and C. K. Yap. A “retraction” method for planning the motion of a disk. *J. Algorithms*, 6:104–111, 1985.
- [PR93] R. Pollack and M.-F. Roy. On the number of cells defined by a set of polynomials. *C. R. Acad. Sci. Paris*, 316:573–577, 1993.
- [PT03] J. Pach and G. Tóth. How many ways can one draw a graph?, 2003. Manuscript, accepted for publication in *Combinatorica*.
- [Rad21] J. Radon. Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten. *Math. Ann.*, 83:113–115, 1921.
- [Rad47] R. Rado. A theorem on general measure. *J. London Math. Soc.*, 21:291–300, 1947.
- [RBG01] L. Rónyai, L. Babai, and M. K. Ganapathy. On the number of zero-patterns of a sequence of polynomials. *J. Amer. Math. Soc.*, 14(3):717–735 (electronic), 2001.
- [RG97] J. Richter-Gebert. *Realization Spaces of Polytopes*. Lecture Notes in Mathematics 1643. Springer, Berlin, 1997.
- [RG99] J. Richter-Gebert. The universality theorems for oriented matroids and polytopes. In B. Chazelle et al., editors, *Advances in Discrete and Computational Geometry*, Contemp. Math. 223, pages 269–292. Amer. Math. Soc., Providence, RI, 1999.
- [Rud91] W. Rudin. *Functional Analysis (2nd edition)*. McGraw-Hill, New York, 1991.
- [Sal75] G.T. Sallee. A Helly-type theorem for widths. In *Geom. Metric Lin. Spaces, Proc. Conf. East Lansing 1974*, Lect. Notes Math. 490, pages 227–232. Springer, Berlin etc., 1975.
- [Sch01] L. Schläfli. Theorie der vielfachen Kontinuität. *Denkschriften der Schweizerischen naturforschender Gesellschaft*, 38:1–237, 1901. Written in 1850–51. Reprinted in *Ludwig Schläfli, 1814–1895, Gesammelte mathematische Abhandlungen*, Birkhäuser, Basel 1950.

- [Sch11] P. H. Schoute. Analytic treatment of the polytopes regularly derived from the regular polytopes. *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam*, 11(3), 1911.
- [Sch86] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience, New York, NY, 1986.
- [Sei91] R. Seidel. Small-dimensional linear programming and convex hulls made easy. *Discrete Comput. Geom.*, 6:423–434, 1991.
- [Sei97] R. Seidel. Convex hull computations. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 19, pages 361–376. CRC Press LLC, Boca Raton, FL, 1997.
- [Sha01] M. Sharir. The Clarkson–Shor technique revisited and extended. In *Proc. 17th Annu. ACM Sympos. Comput. Geom.*, pages 252–256, 2001.
- [Sho91] P. W. Shor. Stretchability of pseudolines is NP-hard. In P. Gritzman and B. Sturmfels, editors, *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift*, volume 4 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 531–554. AMS Press, 1991.
- [Sib81] R. Sibson. A brief description of natural neighbour interpolation. In V. Barnett, editor, *Interpreting Multivariate Data*, pages 21–36. John Wiley & Sons, Chichester, 1981.
- [Ste26] J. Steiner. Einige Gesetze über die Theilung der Ebene und des Raumes. *J. Reine Angew. Math.*, 1:349–364, 1826.
- [Ste16] E. Steinitz. Bedingt konvergente Reihen und konvexe Systeme I; II; III. *J. Reine Angew. Math.*, 143; 144; 146:128–175; 1–40; 1–52, 1913; 1914; 1916.
- [Ste85] H. Steinlein. Borsuk's antipodal theorem and its generalizations and applications: a survey. In A. Granas, editor, *Méthodes topologiques en analyse nonlinéaire*, pages 166–235. Colloq. Sémin. Math. Super., Semin. Sci. OTAN (NATO Advanced Study Institute) 95, Univ. de Montréal Press, Montréal, 1985.
- [SU00] J.-R. Sack and J. Urrutia, editors. *Handbook of Computational Geometry*. North-Holland, Amsterdam, 2000.
- [Tho65] R. Thom. On the homology of real algebraic varieties (in French). In S.S. Cairns, editor, *Differential and Combinatorial Topology*. Princeton Univ. Press, 1965.
- [Vin39] P. Vincensini. Sur une extension d'un théorème de M. J. Radon sur les ensembles de corps convexes. *Bull. Soc. Math. France*, 67:115–119, 1939.
- [Vor08] G. M. Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième Mémoire: Recherches sur les paralléloèdres primitifs. *J. Reine Angew. Math.*, 134:198–287, 1908.

-
- [War68] H. E. Warren. Lower bound for approximation by nonlinear manifolds. *Trans. Amer. Math. Soc.*, 133:167–178, 1968.
- [Wel01] E. Welzl. Entering and leaving j -facets. *Discrete Comput. Geom.*, 25:351–364, 2001.
- [Zas75] T. Zaslavsky. *Facing up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes*, volume 154 of *Memoirs Amer. Math. Soc.* American Mathematical Society, Providence, RI, 1975.
- [Zie94] G. M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 1994. Corrected and revised printing 1998.
- [ŽV90] R.T. Živaljević and S.T. Vrećica. An extension of the ham sandwich theorem. *Bull. London Math. Soc.*, 22:183–186, 1990.