Lecture Notes on Polytopes

Anthony Hong[1]

March 6, 2024

[1]Prof. Laura Escobar Vega's course Math547 (SP24) Topics in Geometry: Theory of Polytopes.

Contents

Chapter 1

Introduction to Polytopes

1.1 Why study polytopes?

- Classical: Euclid's Elements presents the platonic solids as a crowning achievement of Greek mathematics.
- Useful: Linear optimization is equivalent to finding points in polytopes.
- Interdisciplinary: Provide combinatorial tools to other areas of mathematics (e.g. symplectic geometry, algebraic geometry, number theory, etc.)
- Fun for some.

1.2 What is a polytope?

1.2.1 Affine subspaces

The nonempty **affine subspaces**, or **flats**, are the translates of linear subspaces (the vector subspaces of $\mathbb{R}^d$ containing the origin $0 \in \mathbb{R}^d$). The **dimension of an affine subspace** is the dimension of the corresponding linear vector space. Affine subspaces of dimensions $0, 1, 2$, and $d-1$ in $\mathbb{R}^d$ are called points, lines, planes, and hyperplanes, respectively. We take for granted the fact that affine subspaces can be described by as the affine image of some real vector space $\boldsymbol{a} + V$ (where V is a linear subspace) or as the set of all affine combinations of a finite set of points,

$$F = \left\{ \boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \lambda_0 \boldsymbol{x}_0 + \ldots + \lambda_n \boldsymbol{x}_n \text{ for } \lambda_i \in \mathbb{R}, \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

That is, every affine subspace can be described both as an intersection of affine hyperplanes, and as the **affine hull** of a finite point set (i.e., as the intersection of all affine flats that contain the set). A set of $n \geqslant 0$ points is **affinely independent** if its affine hull has dimension $n-1$, that is, if every proper subset has a smaller affine hull.

Proposition 1.2.1. The two definitions of affine subspace $\boldsymbol{a} + V$ and $\{\sum \lambda_i \boldsymbol{x}_i \mid \sum \lambda_i = 1\}$ are equivalent.

Proof.
From $\boldsymbol{a} + V$ to Affine Combinations:
Given $\boldsymbol{a} + V$, where $\boldsymbol{a}$ is a particular point and V is a vector space, any point in $\boldsymbol{a} + V$ can be written as $\boldsymbol{a} + \boldsymbol{v}$, where $\boldsymbol{v} \in V$. If we choose a basis $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ for V, then any $\boldsymbol{v} \in V$ can be expressed as a

linear combination $\boldsymbol{v} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$, where λ_i are scalars. Therefore, any point in $\boldsymbol{a} + V$ can be written as $\boldsymbol{a} + \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$. If we set $\lambda_0 = 1 - \sum_{i=1}^{k} \lambda_i$, then we can write this as $\lambda_0 \boldsymbol{a} + \sum_{i=1}^{k} \lambda_i (\boldsymbol{a} + \boldsymbol{x}_i)$, ensuring that $\sum_{i=0}^{k} \lambda_i = 1$. This shows that every point in $\boldsymbol{a} + V$ can be seen as an affine combination of points in the subspace.

From Affine Combinations to $\boldsymbol{a} + V$:
Conversely, consider a set defined by affine combinations $\{\sum_{i=0}^{n} \lambda_i \boldsymbol{x}_i \mid \sum \lambda_i = 1\}$. Let's choose one of these points, say $\boldsymbol{x}_0$, to play the role of $\boldsymbol{a}$ in the $\boldsymbol{a} + V$ definition. We can then view the differences $\boldsymbol{x}_i - \boldsymbol{x}_0$ as elements of a vector space V, since they represent directions (or displacements) from $\boldsymbol{x}_0$ to other points in the set. This shows that the set of affine combinations can be expressed as $\boldsymbol{a} + V$, where $\boldsymbol{a} = \boldsymbol{x}_0$ and V is the span of $\{\boldsymbol{x}_i - \boldsymbol{x}_0\}$. □

1.2.2 Polytopes

A point set $K \subseteq \mathbb{R}^d$ is **convex** if with any two points $\boldsymbol{x}, \boldsymbol{y} \in K$ it also contains the straight line segment $[\boldsymbol{x}, \boldsymbol{y}] = \{\lambda x + (1-\lambda) y : 0 \leqslant \lambda \leqslant 1\}$.

Clearly, every intersection of convex sets is convex, and $\mathbb{R}^d$ itself is convex. Thus for any $K \subseteq \mathbb{R}^d$, the "smallest" convex set containing K, called the **convex hull** of K, can be constructed as the intersection of all convex sets that contain K :

$$\operatorname{conv}(K) := \bigcap \left\{K' \subseteq \mathbb{R}^d : K \subseteq K', K' \text{ convex } \right\}$$

For any finite set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq K$ and parameters $\lambda_1, \ldots, \lambda_k \geqslant 0$ with $\lambda_1 + \ldots + \lambda_k = 1$, the convex hull $\operatorname{conv}(K)$ must contain the point $\lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k$: this can be seen by induction on k, using

$$\lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k = (1-\lambda_k)\left(\frac{\lambda_1}{1-\lambda_k}\boldsymbol{x}_1 + \ldots + \frac{\lambda_{k-1}}{1-\lambda_k}\boldsymbol{x}_{k-1}\right) + \lambda_k \boldsymbol{x}_k$$

for $\lambda_k < 1$. When $k = 1$, the convex hull of a single point is itself. When $k = 2$, every convex set containing $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ must contain $[\boldsymbol{x}_1, \boldsymbol{x}_2]$, so their intersection has to contain $[\boldsymbol{x}_1, \boldsymbol{x}_2]$. Then do the induction on k, the size of finite subset in K, by above formula. This will show the $\supseteq$ direction of the following relationship:

$$\operatorname{conv}(K) = \left\{\lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k \middle| \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq K, \lambda_i \geqslant 0, \sum_{i=1}^{k} \lambda_i = 1\right\}$$

But the right-hand side of this equation is easily seen to be convex, which proves the equality.

Now if $K = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq \mathbb{R}^d$ is itself finite, then we get the definition of a polytope.

Definition 1.2.2. A **polytope**, or a **$\mathcal{V}$-polytope**, is the convex hull of a finite set of points in some $\mathbb{R}^d$.

$$\operatorname{conv}(K) = \left\{\lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_n \boldsymbol{x}_n : n \geqslant 1, \lambda_i \geqslant 0, \sum_{i=1}^{n} \lambda_i = 1\right\}.$$

We consider a generalization.

Definition 1.2.3. A **cone** is a nonempty set of vectors $C \subseteq \mathbb{R}^d$ that with any finite set of vectors also contains all their linear combinations with nonnegative coefficients. In particular, every cone contains 0 . For an arbitrary subset $Y \subseteq \mathbb{R}^d$, we define its conical hull (or positive hull) cone (Y) as the intersection of all cones in $\mathbb{R}^d$ that contain Y. Clearly $C := \operatorname{cone}(Y)$ is a cone for every Y. Similar to the situation for convex hulls (Lecture 0), one can easily see that

$$\operatorname{cone}(Y) = \{\lambda_{\mathbf{1}} \boldsymbol{y}_1 + \ldots + \lambda_k \boldsymbol{y}_k : \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\} \subseteq Y, \lambda_i \geqslant 0\}$$

In the case where $Y = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\} \subseteq \mathbb{R}^d$ is a finite set - this is the only case we will need here - this reduces to

$$\operatorname{cone}(Y) := \{t_1\boldsymbol{y}_1 + \ldots + t_n\boldsymbol{y}_n : t_i \geqslant 0\} = \{Y\boldsymbol{t} : \boldsymbol{t} \geqslant 0\}$$

We define that $\operatorname{cone}(Y) = \{0\}$ if Y is the empty set, i.e., if $n = 0$. The **vector sum** (or **Minkowski sum**) of two sets $P, Q \subseteq \mathbb{R}^d$ is defined to be

$$P + Q := \{\boldsymbol{x} + \boldsymbol{y} : \boldsymbol{x} \in P, \boldsymbol{y} \in Q\}$$

Definition 1.2.4. A $\mathcal{V}$**-polyhedron** is any finitely generated convexconical combination: a set $P \subseteq \mathbb{R}^d$ that is given in the form

$$P = \operatorname{conv}(V) + \operatorname{cone}(Y) \quad \text{for some } V \in \mathbb{R}^{d\times n}, Y \in \mathbb{R}^{d\times n'},$$

as the Minkowski sum of a convex hull of a finite point set and the cone generated by a finite set of vectors.

Thus, comparing this to definition of a polytope we get that a $\mathcal{V}$-polytope is a $\mathcal{V}$ polyhedron that is bounded, that is, contains no ray $\{\boldsymbol{u} + t\boldsymbol{v} : t \geqslant 0\}$ with $v \neq 0$. For this we only need to observe that $\operatorname{conv}(V)$ is always bounded. This follows from a trivial computation: if $\boldsymbol{x} \in \operatorname{conv}(V)$, then

$$\min\{v_{ik} : 1 \leqslant i \leqslant n\} \leqslant x_k \leqslant \max\{v_{ik} : 1 \leqslant i \leqslant n\},$$

which encloses $\operatorname{conv}(V)$ in a bounded box.

The **dimension** of a polytope is the dimension of its affine hull. A d**-polytope** is a polytope of dimension d in some $\mathbb{R}^e (e \geqslant d)$. Two polytopes $P \subseteq \mathbb{R}^d$ and $Q \subseteq \mathbb{R}^e$ are **affinely isomorphic**, denoted by $P \cong Q$, if there is an affine map $f : \mathbb{R}^d \longrightarrow \mathbb{R}^e$ that is a bijection between the points of the two polytopes. (Note that such a map need not be injective or surjective on the "ambient spaces.")

Example 1.2.5.
The standard d-simplex is $\Delta_d := \operatorname{conv}\{e_1, \ldots, e_{d+1}\} \subseteq \mathbb{R}^{d+1}$
The d-cube is $C_d := \operatorname{conv}\{0,1\}^d = [0,1]^d \subseteq \mathbb{R}^d$. In fact, $C_d = \left\{\boldsymbol{x} \in \mathbb{R}^d \mid 0 \leqslant x_i \leqslant 1\right\}$.
The d-cross polytope is $\diamond_d := \operatorname{conv}\{\pm e_1, \ldots, \pm e_d\} \subseteq \mathbb{R}^d$.
Two-dimensional polytopes are called polygons.

We consider another approach to define polyhedron and polytope.

Definition 1.2.6. An $\mathcal{H}$**-polyhedron** is an intersection of finitely many closed halfspaces in some $\mathbb{R}^d$. An $\mathcal{H}$**-polytope** is an $\mathcal{H}$-polyhedron that is bounded in the sense that it does not contain a ray $\{\boldsymbol{x} + t\boldsymbol{y} : t \geqslant 0\}$ for any $\boldsymbol{y} \neq 0$. An $\mathcal{H}$-polyhedron can be represented by

$$P = P(A, \boldsymbol{z}) = \left\{\boldsymbol{x} \in \mathbb{R}^d : A\boldsymbol{x} \leqslant \boldsymbol{z}\right\} \quad \text{for some } A \in \mathbb{R}^{m\times d}, \boldsymbol{z} \in \mathbb{R}^m.$$

(Here " $A\boldsymbol{x} \leqslant \boldsymbol{z}$ " is the usual shorthand for a system of inequalities, namely $\boldsymbol{a}_1\boldsymbol{x} \leqslant z_1, \ldots, \boldsymbol{a}_m\boldsymbol{x} \leqslant z_m$, where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ are the rows of A, and $z_1, \ldots, z_m$ are the components of $\boldsymbol{z}$.)

We now show that the two definitions are equivalent.

Theorem 1.2.7 (Main theorem for polytopes)**.**

$$\{\mathcal{H}\text{-polytope}\} = \{\mathcal{V}\text{-polytope}\}.$$

A subset $P \subseteq \mathbb{R}^d$ is the convex hull of a finite point set (a $\mathcal{V}$-polytope)

$$P = \operatorname{conv}(V) \quad \text{for some } V \in \mathbb{R}^{d\times n}$$

if and only if it is a bounded intersection of halfspaces (an $\mathcal{H}$-polytope)

$$P = P(A, z) \quad \text{for some } A \in \mathbb{R}^{m\times d}, z \in \mathbb{R}^m$$

This result contains two implications, which are equally "geometrically clear" and nontrivial to prove, and which in a certain sense are equivalent.

This theorem provides two independent characterizations of polytopes that are of different power, depending on the problem we are studying. For example, consider the following four statements.
- Every intersection of a polytope with an affine subspace is a polytope.
- Every intersection of a polytope with a polyhedron is a polytope.
- The Minkowski sum of two polytopes is a polytope.
- Every projection of a polytope is a polytope.

The first two statements are trivial for a polytope presented in the form $P = P(A, \boldsymbol{z})$ (where the first is a special case of the second), but both are nontrivial for the convex hull of a finite set of points. Similarly the last two statements are easy to see for the convex hull of a finite point set, but are nontrivial for bounded intersections of halfspaces.

Theorem 1.2.7 is the version we really need, a very basic statement about polytopes; however, it is not the most straightforward version to prove. Therefore we generalize it to a theorem about polyhedra, due to Motzkin.

Theorem 1.2.8 (Main theorem for polyhedra)**.**

$$\{\mathcal{H}\text{-polyhedron} = \mathcal{V}\text{-polyhedron}\}.$$

A subset $P \subseteq \mathbb{R}^d$ is a sum of a convex hull of a finite set of points plus a conical combination of vectors (a $\mathcal{V}$-polyhedron)

$$P = \operatorname{conv}(V) + \operatorname{cone}(Y) \quad \text{for some } V \in \mathbb{R}^{d\times n}, Y \in \mathbb{R}^{d\times n'}$$

if and only if is an intersection of closed halfspaces (an $\mathcal{H}$-polyhedron)

$$P = P(A, z) \quad \text{for some } A \in \mathbb{R}^{m\times d}, z \in \mathbb{R}^m.$$

First note that Theorem 1.2.7 follows from Theorem 1.2.8: we have already seen that polytopes are bounded polyhedra, in both the $\mathcal{V}$ and the $\mathcal{H}$ versions.

Proof. Sketch of proof of $\supseteq$. Let $P = \operatorname{conv}(V) + \operatorname{cone}(Y)$ and identify V with the $d \times n$ with columns the elements of V and similarly Y with an $d \times m$ matrix. Note that

$$P = \left\{x \in \mathbb{R}^d \,|\exists \lambda \in \mathbb{R}^n, \mu \in \mathbb{R}^m |\, x = V\lambda + Y\mu, \sum \lambda_i = 1, \lambda_i \geqslant 0, \mu_i \geqslant 0\right\}.$$

Let

$$Q = \left\{ \begin{bmatrix} x \\ \lambda \\ \mu \end{bmatrix} \in \mathbb{R}^{d+n+m} \,\middle|\, x \in P \right\}.$$

Note that Q is given by the half-spaces $x - V\lambda - Y\mu \geqslant 0, x - V\lambda - Y\mu \leqslant 0, \sum \lambda_i = 1$, $\lambda_i \geqslant 0, \mu_i \geqslant 0$. Moreover, P is a projection of Q. Thus this direction relies on showing that the projection of an $\mathcal{H}$-polyhedron is an $\mathcal{H}$-polyhedron. This is done using Fourier-Motzkin elimination. □

Example 1.2.9. Suppose Q is the polyhedron given by

$$x_1 - x_2 \leqslant -1, \quad x_1 + x_2 \leqslant 5, \quad -x_1 + x_2 \leqslant 3, \quad -x_1 \leqslant 0$$

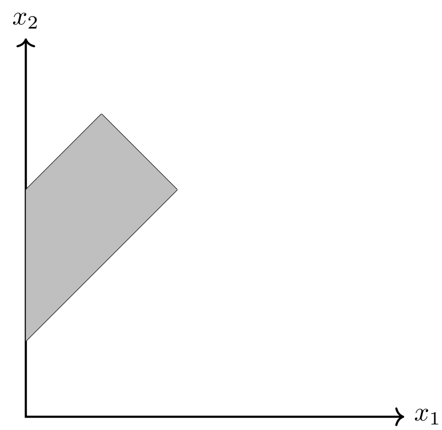
and we wish to project onto the x_1-axis. To do so we should eliminate the x_2-variable. Note,

$$x_1 + 1 \leqslant x_2 \leqslant -x_1 + 5, x_1 + 3.$$

Thus, the projection is given by

$$-x_1 \leqslant 0, x_1 + 1 \leqslant -x_1 + 5, x_1 + 1 \leqslant x_1 + 3$$

which becomes $0 \leqslant x_1 \leqslant 2$. Fourier-Motzkin elimination generalizes this.



Proof. Sketch of proof of $\subseteq$. Let $P = P(A, z)$. Consider $Q = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{d+n} \middle| Ax \leqslant y \right\}$. We will show that Q is a V-polyhedron. Note that $P = Q \cap \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{d+n} \middle| y = z \right\}$, where the latter is an affine hyperplane. We will also show that the intersection of a V-polyhedron with an affine hyperplane is a V-polyhedron.
(1) Q is a $\mathcal{V}$-polyhedron. Note that

$$\begin{aligned} Q &= \left\{ \begin{bmatrix} x \\ Ax + w \end{bmatrix} \middle| x \in \mathbb{R}^d, w \in R^n_{\geqslant 0} \right\} \\ &= \text{cone} \left\{ \begin{bmatrix} \pm e_1 \\ \pm Ae_1 \end{bmatrix}, \ldots, \begin{bmatrix} \pm e_d \\ \pm Ae_d \end{bmatrix}, \begin{bmatrix} 0 \\ f_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ f_n \end{bmatrix} \right\} \end{aligned}$$

where $e_1, \ldots, e_d \in \mathbb{R}^d \times 0$ are the standard basis vectors and $f_1, \ldots, f_n \in 0 \times \mathbb{R}^n$ are the standard basis vectors.
(2) The intersection of a $\mathcal{V}$-polyhedron with an affine hyperplane is a $\mathcal{V}$-polyhedron. We are skipping the proof. □

1.3 Farkas Lemma

The following version of Farkas lemma yields a characterization for the solvability of a system of inequalities (we are using [1]'s numbering).

Proposition 1.3.1 (Farkas lemma I). Let $A \in \mathbb{R}^{m \times d}$ and $z \in \mathbb{R}^m$. Either
(i) there exists a point $\boldsymbol{x} \in \mathbb{R}^d$ with $A\boldsymbol{x} \leqslant \boldsymbol{z}$, or
(ii) there exists a row vector $\boldsymbol{c} \in (\mathbb{R}^m)^*$ with $\boldsymbol{c} \geqslant \mathbb{0}, \boldsymbol{c}A = \mathbb{0}$ and $\boldsymbol{cz} < 0$,
but not both.

The next version of Farkas Lemma states that either a system of equations has a positive solution or a vector that certifies that such a solution does not exist.

Proposition 1.3.2 (Farkas lemma II). Let $A \in \mathbb{R}^{m \times d}$ and $\boldsymbol{z} \in \mathbb{R}^m$. Either
(i) there exists a point $\boldsymbol{x} \in \mathbb{R}^d$ with $A\boldsymbol{x} = \boldsymbol{z}, \boldsymbol{x} \geqslant 0$, or
(ii) there exists a row vector $\boldsymbol{c} \in (\mathbb{R}^m)^*$ with $\boldsymbol{c}A \geqslant \mathbb{0}$ and $\boldsymbol{cz} < 0$,
but not both.

Proof. We have the following equivalences:

$$
\begin{aligned}
&\exists \boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{z}, \boldsymbol{x} \geqslant \mathbf{0} \\
&\Longleftrightarrow \exists \boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{z}, (-A)\boldsymbol{x} \leqslant -\boldsymbol{z}, -\boldsymbol{x} \leqslant \mathbf{0} \\
&\Longleftrightarrow \exists \boldsymbol{x} : \begin{pmatrix} A \\ -A \\ -I_d \end{pmatrix} \boldsymbol{x} \leqslant \begin{pmatrix} \boldsymbol{z} \\ -\boldsymbol{z} \\ \mathbf{0} \end{pmatrix} \\
&\overset{\text{FL I}}{\Longleftrightarrow} \nexists c_1 \geqslant \mathbb{0}, c_2 \geqslant \mathbb{0}, b \geqslant 0 : \\
&\qquad (\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{b}) \begin{pmatrix} A \\ -A \\ -I_d \end{pmatrix} = \mathbb{0}, (\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{b}) \begin{pmatrix} \boldsymbol{z} \\ -\boldsymbol{z} \\ \mathbf{0} \end{pmatrix} < 0 \\
&\Longleftrightarrow \nexists \boldsymbol{c}_1 \geqslant \mathbb{0}, \boldsymbol{c}_2 \geqslant \mathbb{0}, \boldsymbol{b} \geqslant \mathbb{0} : (\boldsymbol{c}_1 - \boldsymbol{c}_2) A - \boldsymbol{b} = \mathbb{0}, (\boldsymbol{c}_1 - \boldsymbol{c}_2) \boldsymbol{z} < 0 \\
&\Longleftrightarrow \nexists \boldsymbol{c} = \boldsymbol{c}_1 - \boldsymbol{c}_2, \boldsymbol{b} \geqslant \mathbb{0} : \boldsymbol{c}A - \boldsymbol{b} = \mathbb{0}, \boldsymbol{c}\boldsymbol{z} < 0 \\
&\Longleftrightarrow \nexists \boldsymbol{c} : \boldsymbol{c}A \geqslant \mathbb{0}, \boldsymbol{c}\boldsymbol{z} < 0.
\end{aligned}
$$

□

Proposition 1.3.3 (Farkas lemma IV)**.** Let $V \in \mathbb{R}^{d \times n}, Y \in \mathbb{R}^{d \times n'}$, and $\boldsymbol{x} \in \mathbb{R}^d$. Either
(i) there exist $\boldsymbol{t}, \boldsymbol{u} \geqslant \mathbf{0}$ with $\mathbb{1}\boldsymbol{t} = \mathbf{1}$ and $\boldsymbol{x} = V\boldsymbol{t} + \boldsymbol{Y}\boldsymbol{u}$, or
(ii) there exists a row vector $(\alpha, \boldsymbol{a}) \in \left(\mathbb{R}^{d+1}\right)^*$ with $\boldsymbol{a}\boldsymbol{v}_i \leqslant \alpha$ for all $i \leqslant n$, $\boldsymbol{a}\boldsymbol{y}_j \leqslant 0$ for all $j \leqslant n'$, while $\boldsymbol{a}\boldsymbol{x} > \alpha$, but not both.

Proof. The "either" condition can be stated as

$$
\exists \begin{pmatrix} \boldsymbol{t} \\ \boldsymbol{u} \end{pmatrix} \geqslant \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} : \quad \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ V & Y \end{pmatrix} \begin{pmatrix} \boldsymbol{t} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}
$$

which by version II of the Farkas lemma is equivalent to

$$
\begin{aligned}
&\overset{\text{FL II}}{\Longleftrightarrow} \nexists (\alpha, -\boldsymbol{a}) \in \left(\mathbb{R}^{d+1}\right)^* : \quad (\alpha, -\boldsymbol{a}) \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ V & Y \end{pmatrix} \geqslant (\mathbb{0}, \mathbb{0}),\ (\alpha, -a) \begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} < 0 \\
&\Longleftrightarrow \nexists (\alpha, -\boldsymbol{a}) \in \left(\mathbb{R}^{d+\mathbf{1}}\right)^* : \quad \alpha\mathbb{1} - \boldsymbol{a}V \geqslant \mathbb{0}, \boldsymbol{a}Y \leqslant \mathbb{0}, \boldsymbol{a}\boldsymbol{x} > \alpha,
\end{aligned}
$$

which is equivalent to the negation of the "or" condition. □

1.4 Faces of polytopes

Definition 1.4.1. Let $P \subseteq \mathbb{R}^d$ be a convex polytope. Let $\boldsymbol{c}$ be a row vector. A linear inequality $\boldsymbol{c}\boldsymbol{x} \leqslant c_0$ is **valid** for P if it is satisfied for all points $\boldsymbol{x} \in P$. A **face** of P is any set of the form

$$
F = P \cap \left\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{c}\boldsymbol{x} = c_0\right\}
$$

where $\boldsymbol{c}\boldsymbol{x} \leqslant c_0$ is a valid inequality for P. Thus, equivalently, if $\boldsymbol{c}$ is a column vector, a **face** of P can also be written as

$$
F = \{\boldsymbol{x} \in P : \forall \boldsymbol{y} \in P, \langle \boldsymbol{y}, \boldsymbol{c} \rangle \leqslant \langle \boldsymbol{x}, \boldsymbol{c} \rangle\}
$$

The dimension of a face is the **dimension** of its affine hull: $\dim(F) := \dim(\operatorname{aff}(F))$.

For the valid inequality $\mathbb{0}\mathbf{x} \leqslant 0$, we get that P itself is a face of P. All other faces of P, satisfying $F \subset P$, are called **proper faces**. For the inequality $\mathbb{0}\mathbf{x} \leqslant 1$, we see that $\varnothing$ is always a face of P. The faces of dimensions $0, 1, \dim(P)-2$, and $\dim(P)-1$ are called **vertices**, **edges**, **ridges**, and **facets**, respectively. Thus, in particular, the vertices are the minimal nonempty faces, and the facets are the maximal proper faces. The set of all vertices of P, the vertex set, will be denoted by vert (P).

Example 1.4.2. Let $P = \operatorname{conv}(0, e_1, e_2)$.
- If $m = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, then $P_m = \{0\}$.
- If $m = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then $P_m = \operatorname{conv}(e_1, e_2)$.
- If $m = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $P_m = \{e_2\}$.
- If $m = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then $P_m = P$.

Proposition 1.4.3 (ziegler proposition 2.2)**.** Let $P \subseteq \mathbb{R}^d$ be a polytope.
(i) Every polytope is the convex hull of its vertices: $P = \operatorname{conv}(\operatorname{vert}(P))$.
(ii) If a polytope can be written as the convex hull of a finite point set, then the set contains all the vertices of the polytope: $P = \operatorname{conv}(V)$ implies that $\operatorname{vert}(P) \subseteq V$.

Proof. Write $P = \operatorname{conv}(V)$ with V finite. If any $v \in V$ can be written as a convex combination of elements in $V' := V - \{v\}$ then $P = \operatorname{conv}(V')$. Repeat until no longer possible until we get $P = \operatorname{conv}(W)$. We claim that $W = \operatorname{verts}(P)$.

$\supseteq$: Let $v = \lambda_1 w_1 + \cdots + \lambda_n w_n \in \operatorname{verts}(P)$ with $w_1, \ldots, w_n \in W, \sum \lambda_i = 1$, and $\lambda_i \geqslant 0$. Let c be such that $P_c = \{v\}$ and note that for all $i, \langle c, w_i\rangle \langle\langle c, v\rangle$. It follows that

$$\langle c, v\rangle = \sum \lambda_i \langle c, w_i\rangle < \langle c, v\rangle$$

a contradiction.

$\subseteq$: Let $w \in W$ and consider $W' = W - \{w\}$. Since $w \notin \operatorname{conv}(W')$ there does not exist $t \geqslant 0$ such that $w = W't$ and $\mathbb{1}t = 1$. Equivalently, there does not exist $t \geqslant 0$ such that

$$\begin{bmatrix} \mathbb{1} \\ W' \end{bmatrix} t = \begin{bmatrix} 1 \\ w \end{bmatrix}.$$

By Farkas Lemma II, there exists c such that $c \begin{bmatrix} \mathbb{1} \\ W' \end{bmatrix} \geqslant 0$ and $c \begin{bmatrix} 1 \\ w \end{bmatrix} < 0$. Writing $c = (\beta, -b)$, then $\beta \mathbb{1} - bW' \geqslant 0$ and $\beta - bw < 0$. It follows that $bW' \leqslant (\beta, \ldots, \beta)$ and $bw > \beta$, i.e. $P_b = \{w\}$. ☐

Proposition 1.4.4 (ziegler proposition 2.3)**.** Let $P \subseteq \mathbb{R}^d$ be a polytope, and $V := \operatorname{vert}(P)$. Let F be a face of P.
(i) The face F is a polytope, with $\operatorname{vert}(F) = F \cap V$.
(ii) Every intersection of faces of P is a face of P.
(iii) The faces of F are exactly the faces of P that are contained in F.
(iv) $F = P \cap \operatorname{aff}(F)$.

We will need another construction: the **vertex figure** obtained by cutting a polytope by a hyperplane that cuts off a single vertex.

For this, we consider a polytope P with $V = \operatorname{vert}(P)$, and a vertex $v \in V$. Let $\boldsymbol{c}\boldsymbol{x} \leqslant c_0$ be a valid inequality with

$$\{\boldsymbol{v}\} = P \cap \{\boldsymbol{x} : \boldsymbol{c}\boldsymbol{x} = c_0\}.$$

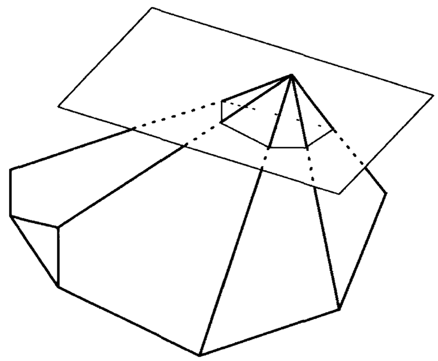


Figure 1.1: Vertex figure.

Furthermore, we choose some $c_1 < c_0$ with $\boldsymbol{c}\boldsymbol{v}' < c_1$ for all $\boldsymbol{v}' \in \operatorname{vert}(P) \backslash \boldsymbol{v}$. Then we define a vertex figure of P at v as the polytope

$$P/\boldsymbol{v} := P \cap \{\boldsymbol{x} : \boldsymbol{c}\boldsymbol{x} = c_1\}.$$

Note that the construction of P/v depends on the choice of c_1 and of the inequality $\boldsymbol{c}\boldsymbol{x} \leqslant c_0$; however, the following result shows that the combinatorial type of P/v is independent of this.

Proposition 1.4.5 (ziegler Proposition 2.4)**.** There is a bijection between the k-dimensional faces of P that contain v, and the $(k-1)$-dimensional faces of P/v, given by

$$\begin{aligned} \pi : \quad F &\longmapsto F \cap \{\boldsymbol{x} : \boldsymbol{c}\boldsymbol{x} = c_1\}, \\ \sigma : P \cap \operatorname{aff}\big(\{v\} \cup F'\big) &\longleftarrow F'. \end{aligned}$$

1.4.1 Face lattices

A **partial ordering** on a (nonempty) set S is a binary relation on S, denoted $\leqslant$, which satisfies the following properties:

- reflexive: for all $s \in S, s \leqslant s$,
- antisymmetric: if $s \leqslant s'$ and $s' \leqslant s$ then $s = s'$,
- transitive: if $s \leqslant s'$ and $s' \leqslant s''$ then $s \leqslant s''$.

When we fix a partial ordering $\leqslant$ on S, we refer to S (or, more precisely, to the pair $(S, \leqslant)$) as a **partially ordered set**, also abbreviated as **poset**.

It is important to notice that we do not assume all pairs of elements in S are **comparable** under $\leqslant$: for some s and s' we may have neither $s \leqslant s'$ nor $s' \leqslant s$. If all pairs of elements can be compared (that is, for all s and s' in S either $s \leqslant s'$ or $s' \leqslant s$) then we say S is **totally ordered** with respect to $\leqslant$.

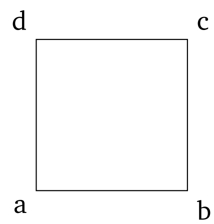
A **chain** in S is a totally ordered subset of S; its **length** is its number of elements minus 1.

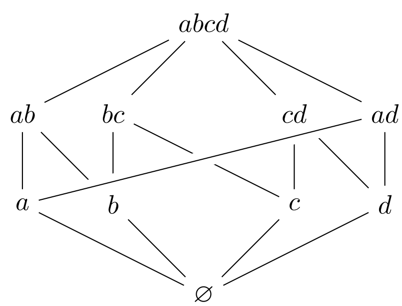
Example 1.4.6. The usual ordering relation $\leqslant$ on $\mathbb{R}$ or on $\mathbb{Z}^+$ is a partial ordering of these sets. In fact it is a total ordering on either set. This ordering on $\mathbb{Z}^+$ is the basis for proofs by induction.

Example 1.4.7. On $\mathbb{Z}^+$, declare $a \leqslant b$ if $a \mid b$. This partial ordering on $\mathbb{Z}^+$ is different from the one in previous example and is called ordering by divisibility. It is one of the central relations in number theory. (Proofs about $\mathbb{Z}^+$ in number theory sometimes work not by induction, but by starting on primes, then extending to prime powers, and then extending to all positive integers using prime factorization. Such proofs view $\mathbb{Z}^+$ through the divisibility relation rather than through the usual ordering relation.) Unlike the ordering on $\mathbb{Z}^+$ in previous example, $\mathbb{Z}^+$ is not totally ordered by divisibility: most pairs of integers are not comparable under the divisibility relation. For instance, 3 doesn't divide 5 and 5 doesn't divide 3 . The subset $\{1, 2, 4, 8, 16, \ldots\}$ of powers of 2 is totally ordered under divisibility.

Definition 1.4.8. The **face lattice** of a convex polytope P is the poset $L := L(P)$ of all faces of P, partially ordered by inclusion.

Example 1.4.9.
(1) The **Boolean lattice** is the poset given by $(2^{[d]}, \subseteq)$, where we use $[d]$ to denote $\{1, \cdots, d\}$ and 2^X is the power set of X.
(2) Face lattice $L(C_2)$ of cycle C_2 using **Hasse's diagram** of poset (the element in the poset that is higher contains those that are lower). The top is the whole polytope, and the bottom is the empty face. The second line has the edges, and the third line has the vertices.





(3) Exercise: show that $L(\triangle_d)$ is the Boolean lattice.

For elements $x, y \in S$ with $x \leqslant y$, we denote by

$$[x, y] := \{w \in S : x \leqslant w \leqslant y\}$$

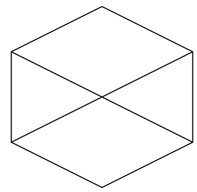
the **interval** between x and y. An interval in S is **boolean** if it is isomorphic to the poset $B_k = \left(2^{[k]}, \subseteq\right)$ of all subsets of a k-element set, for some k.

A poset is **bounded** if it has a unique minimal element, denoted $\hat{0}$, and a unique maximal element, denoted $\hat{1}$. The **proper part** of a bounded poset S is $\bar{S} := S \backslash \{\hat{0}, \hat{1}\}$.

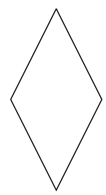
A poset is **graded** if it is bounded, and every maximal chain has the same length. In this case the length of a maximal chain in the interval $[\hat{0}, x]$ is the **rank** of x, denoted by $r(x)$. The rank $r(S) := r(\hat{1})$ is also called the **length** of S. For example, every chain is a graded poset, with $r(C) = |C| - 1$, and the boolean posets B_k are graded of length $r\left(B_k\right) = k$, for all $k \geqslant -1$.

A poset is a **lattice** if it is bounded, and every two elements $x, y \in S$ have a unique minimal upper bound in S, called the **join** $x \vee y$, and every two elements $x, y \in S$ have a unique maximal lower bound in S, called the **meet** $x \wedge y$. (In fact, any two of these three conditions imply the third; also, if every pair of elements has a join respectively meet, then also every finite subset has a join respectively meet.)

Example 1.4.10. The following poset is not a lattice.



Theorem 1.4.11 (ziegler Theorem 2.7.)**.** Let P be a convex polytope.
(i) For every polytope P the face poset $L(P)$ is a graded lattice of length $\dim(P)+1$, with rank function $r(F)=\dim(F)+1$.
(ii) Every interval $[G, F]$ of $L(P)$ is the face lattice of a convex polytope of dimension $r(F)-r(G)-1$.
(iii) ("Diamond property") Every interval of length 2 has exactly four elements. That is, if $G \subseteq F$ with $r(F)-r(G)=2$, then there are exactly two faces H with $G \subset H \subset F$, and the interval $[G, F]$ looks like



Proof. To see that $L(P)$ is a lattice it suffices to see that it has a unique maximal element $\hat{1}=P$ and a unique minimal element $\hat{0}=\varnothing$, and that meets exist, with $F \wedge G=F \cap G$; this is true because $F \cap G$ is a face of F and of G, and thus of P, by Proposition 1.4.4(ii). And clearly every face of P that is contained in F and in G must be contained in $F \cap G$.

We continue with part (ii). For this we can assume that $F=P$, by Proposition 1.4.4(iii). Now if $G=\varnothing$, then everything is clear. If $G \neq \varnothing$, then it has a vertex $v \in G$ by Proposition 1.4.3(i), which is a vertex of P by Proposition 1.4.4(iii). Now the face lattice of P / v is isomorphic to the interval $[\{v\}, P]$ of the face lattice $L(P)$, by Proposition 2.4. Thus we are done by induction on $\dim(G)$.

For part (i) it remains to see that the lattice $L(P)$ is graded. If $G \subset F$ are faces of P, then from $G= P \cap \operatorname{aff}(G) \subseteq P \cap \operatorname{aff}(F)=F$, which holds by Proposition 1.4.4(iv), we can conclude that $\operatorname{aff}(G) \subset \operatorname{aff}(F)$, and thus that $\dim(G)<\dim(F)$. So it suffices to show that if $\dim(F)-\dim(G) \geqslant 2$, then there is a face $H \in L(P)$ with $G \subset H \subset F$. But by part (ii) the interval $[G, F]$ is the face lattice of a polytope of dimension at least 1 , so it has a vertex, which yields the desired H. Part (iii) is a special case of (ii): the "diamond" is the face lattice of a 1-dimensional polytope. $\square$

Definition 1.4.12. Two polytopes P, Q are **combinatorially equivalent** if $L(P) \simeq L(Q)$.

Example 1.4.13. Up to combinatorial equivalence, for each n there is exactly one polygon with n vertices.

Recall that last time we also defined the f-vector of P to be $f(P):=(f_{-1}, f_0, \ldots, f_d)$, where f_i is the number of faces of dimension i, and the f-polynomial of P to be $f_P(t):=\sum_{i=0}^d f_i t^i$

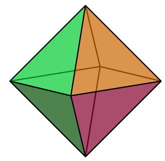
Exercise 1.4.14. Do there exist two non-combinatorially equivalent 3 -dimensional polytopes with the same f-vector?

1.4.2 f-vectors

Definition 1.4.15. The **f-vector** of P is $f(P):=(f_{-1}, f_0, \ldots, f_d)$, where f_i is the number of faces of dimension i. The **f-polynomial** of P is $f_P(t):=\sum_{i=0}^d f_i t^i$.

Example 1.4.16. Let $P=\operatorname{conv}(0, e_1, e_2)$. Then $f(P)=(1,3,3,1)$ and $f_P(t)=3+3t+t^2$.

Example 1.4.17. Let P be an octahedron. Then $f(P)=(1,6,12,8,1)$ and $f_P(t)=6+12t+8t^2+t^3$.



Example 1.4.18. Consider the d-cube $C_d = \{x \in \mathbb{R}^d \mid \forall i, -1 \leqslant x_i \leqslant 1\}$. Given $v \in \mathbb{R}^d$, we have that

$$(C_d)_v = \{x \in C_d \mid v_1 x_1 + \cdots v_d x_d \max\}.$$

Note that - If $v_i > 0$ then $x_i = 1$ maximizes $v_i x_i$.
- If $v_i < 0$ then $x_i = -1$ maximizes $v_i x_i$.
- If $v_i = 0$ then x_i can be anything.

For example, if $v = (+, -, 0, 0, -, +, 0, +, 0)$, then

$$(C_d)_v = \{(1, -1, a, b, -1, 1, c, 1, d) \mid a, b, c, d \in [-1, 1]\} \simeq C_4.$$

It follows that the faces of C_d are in one-to-one correspondence with d-tuples in $\{\pm 1, 0\}^d$. Moreover, the dimension of the face corresponding to a tuple is the number of 0 s. Therefore, $f_k = \left(\begin{array}{c} d \\ k \end{array}\right) 2^{d-k}$ and $f_{C_d}(t) = \sum_0^d \left(\begin{array}{c} d \\ k \end{array}\right) 2^{d-k} t^k = (2+t)^d$.

Exercise 1.4.19. Compute f_{Δ_d}.

A key question in combinatorics asks the following:

Q: What is the structure of the collection of f-vectors of d-dimensional polytopes? (It is also interesting for other manifolds.)

Example 1.4.20. For 2-dimensional polytopes, aka polygons, the answer is simple. The f-vector is $(1, n, n, 1)$ for some n. For 3-dimensional polytopes, the answer is more complicated, but settled.

Theorem 1.4.21 (Euler)**.** Let P be a 3-dimensional polytope. Then

$$f_0 - f_1 + f_2 = 2.$$

Theorem 1.4.22 (Steinitz)**.**

$$\begin{aligned} &\left\{f \in \mathbb{Z}^5 \mid \exists P, f(P) = f\right\} \\ =&\left\{f \in \mathbb{Z}^5 \mid f_{-1} = f_3 = 1, f_0 - f_1 + f_2 = 2, f_2 \leqslant 2f_0 - 4, f_0 \leqslant 2f_2 - 4\right\}. \end{aligned}$$

In arbitrary dimension, much less is known.

Theorem 1.4.23 (Euler-Poincaré equation)**.** Let P be a d-dimensional polytope. Then $-f_{-1} + f_0 + \cdots + (-1)^d f_d = 0$.

Definition 1.4.24. A polytope is **simplicial** if all of its faces are combinatorially equivalent to standard simplices.

Remark 1.4.25. Billera-Lee and Stanley proved the g-Theorem which gives a characterization for the f-vector of simplicial polytopes. This could be a good topic for the long presentation. There are still contributions being done.

Theorem 1.4.26 ((Xue, 20+))**.** Let P be a d-dimensional polytope with $d+s$ vertices, where $s \geqslant 2$ and $d \geqslant s$. Then for every k, $f_k(P) \geqslant \left(\begin{array}{c} d+1 \\ k+1 \end{array}\right) + \left(\begin{array}{c} d \\ k+1 \end{array}\right) - \left(\begin{array}{c} d+1-s \\ k+1 \end{array}\right)$.

Remark 1.4.27 ((Kalais' 3^d conjecture, '89)). If P is centrally symmetric (i.e. v is a vertex if and only if $-v$ is a vertex), then P has at least 3^d nonempty faces, where $d = \dim(P)$.

Remark 1.4.28 (Open question). Is $\left(1, 10^3, 10^5, 10^5, 10^3, 1\right)$ the f-vector of a 4-dimensional polytope?

1.5 Simplicial, Cyclic, and Simple Polytopes

We say that $d+1$ vectors are **affinely independent** if the smallest affine space containing them has dimension d. If P is the convex hull of $d+1$ affinely independent vectors, then P is a d-dimensional polytope and all these vectors are vertices. A d-simplex is the convex hull of $d+1$ affinely independent vectors.

Definition 1.5.1. A d-dimensional polytope is **simplicial** if all of its facets are $(d-1)$ simplices. One can recognize affinely independent vectors by looking at determinants.

Lemma 1.5.2. Let $a_0, \ldots, a_d \in \mathbb{R}^d$. Then $a_0, \ldots, a_d$ are affinely independent if and only if

$$\det \left[\begin{array}{ccc} 1 & \cdots & 1 \\ a_0 & \cdots & a_d \end{array}\right] \neq 0.$$

Definition 1.5.3. Let $d \in \mathbb{N}$. The **moment curve** in $\mathbb{R}^d$ is

$$\mu_d : \mathbb{R} \rightarrow \mathbb{R}^d, \quad t \mapsto \left[t, t^2, \ldots, t^d\right].$$

The **cyclic polytope** $C_d\left(t_1, \ldots, t_n\right) = \operatorname{conv}\left\{\mu_d\left(t_1\right), \ldots, \mu_d\left(t_n\right)\right\}$, where $t_1 < \cdots < t_n$ and $n > d$.

The next theorem says that cyclic polytopes have the largest possible number of faces among all convex polytopes with a given dimension and number of vertices.

Theorem 1.5.4 (Upper bound Theorem - McMullen). For any polytope P of dimension d and n verices we have that $f_k(P) \leqslant f_k\left(C_d\left(t_1, \ldots, t_n\right)\right)$ for any k.

Theorem 1.5.5. Let $d \geqslant 2$.
(1) The cyclic polytope $C_d\left(t_1, \ldots, t_n\right)$ is simplicial.
(2) For $S \subseteq [n]$ with $|S| = d$ we have that $\left\{\mu_d\left(t_s\right) \mid s \in S\right\}$ forms a facet if and only if for all $i < j$ not in S, $|\{k \mid k \in S, i < k < j\}|$ is even.

Lemma 1.5.6 (Vandermonde determinant). Let $a_0, \ldots, a_d \in \mathbb{R}$. Then

$$\det \left[\begin{array}{ccc} 1 & \cdots & 1 \\ a_0 & \cdots & a_d \\ \vdots & & \vdots \\ a_0^d & \cdots & a_d^d \end{array}\right] = \prod_{0 \leqslant i < j \leqslant d} \left(a_j - a_i\right).$$

proof of the theorem. (1) By the lemma, any $d+1$ points $\mu_d\left(t_{i_0}\right), \ldots, \mu_d\left(t_{i_d}\right)$ are affinely independent. It follows that all the $\mu_d\left(t_i\right)$ are vertices and that all the facets are simplices.

(2) Let $S \subseteq [n]$ with $S = \left\{s_1, \ldots, s_d\right\}$. Let H_S be the hyperplane through $\mu_d\left(t_{s_1}\right), \ldots, \mu_d\left(t_{s_d}\right)$. Observe that

$$H_S = \left\{x \in \mathbb{R}^d \,\middle|\, \det \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ x & \mu_d\left(t_{s_1}\right) & \cdots & \mu_d\left(t_{s_d}\right) \end{array}\right] = 0\right\}.$$

Let F_S be the defining equation of H_S. As we can see in Figure 1.5, H_S is a facet if and only if $F_S\left(\mu_d\left(t_i\right)\right)$ has the same sign for all $i \notin S$. The sign of $F_S\left(\mu_d(t)\right)$ changes sign as it passes through its zeroes, which are precisely the $\mu_d\left(t_{s_i}\right)$ (since F_S is a polynomial of degree d). It follows that for $i < j$, $F_S\left(\mu_d\left(t_i\right)\right)$ has the same sign as $F_S\left(\mu_d\left(t_{s_j}\right)\right)$ if and only if there is an even number of sign changes between them.

□

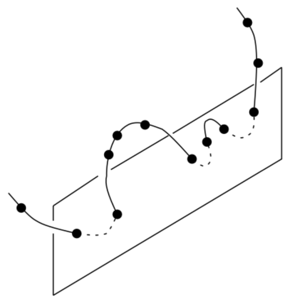


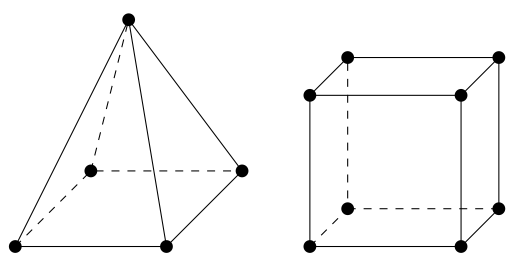
Figure 1.2: H_S with moment curve.

Corollary 1.5.7. The combinatorial type of $C_d\left(t_1, \ldots, t_n\right)$ only depends on d, n.

Sketch of proof. Fix d, n. The V-description of the facets of any cyclic polytope is the same. Since the faces are the intersections of the facets, the structure of the face lattice is the same.

Definition 1.5.8. A polytope of dimension d is **simple** if each vertex is contained in exactly d facets.

Example 1.5.9.
(1) The d-cube is simple.
(2) The pyramid with square base is not simple.



Exercise 1.5.10. Show that if P is simple, then every interval $[F, G]$ of $L(P)$ with $f \neq \varnothing$ is a Boolean lattice.

For these polytopes there is a more compact vector encoding the f-vector.

Definition 1.5.11. Let $f(P) = (f_{-1}, f_0, \ldots, f_d)$. The h-polynomial of P is $h_P(t) = f_P(t-1)$. The h-vector of P is $h(P) = (h_0, \ldots, h_d)$ consisting of the coefficients of h_P.

Example 1.5.12.
(1) Let P be the octahedron. We saw that $f_P(t) = 6 + 12t + 8t^2 + t^3$. It follows that $h_P(t) = 1 - t + 5t^2 + t^3$.
(2) Let P be the 3-cube. Then $f_P(t) = 8 + 12t + 6t^2 + t^3$ and $h_P(t) = 1 + 3t + 3t^2 + t^3$.
(3) We saw $f_{C_d}(t) = (2+t)^d$. It follows that $h_{C_d}(t) = (1+t)^d$.

Theorem 1.5.13 (Dehn-Sommerville equations)**.** Let P be a simple d-polytope. Then $h_i = h_{d-i}$ for all i, i.e. $h_P(t)$ is palindromic.

Reason: if P is simple, then the h-polynomial is the Poincare polynomial of a smooth toric variety. This equations reflect Poincaré duality. There is a method to compute the h-vector of a simple polytope.
(1) Find $\lambda : \mathbb{R}^n \to \mathbb{R}$ linear such that $\lambda(u) \neq \lambda(v)$ for all edges $[u, v]$ of P.
(2) For each vertex u, define $\beta(u) :=\mid \{v \in \operatorname{verts}(P) \mid [u, v] \text{ is an edge}, \lambda(v) > \lambda(u)\} \mid$.

Theorem 1.5.14. Let P be a simple polytope. Then $h_P(t) = \sum_{u \in \operatorname{verts}(P)} t^{\beta(u)}$.

Example 1.5.15. Let $P = \operatorname{conv}\{(w_1, w_2, w_3) \mid w \text{ is a permutation of } [3]\}$. Let $\lambda(x, y, z) = x + 10^2 y + 10^5 z$. The following figure shows the orientation on the edges of P given by $u \to v$ if $\lambda(v) > \lambda(u)$. It follows that $h_P(t) = 1 + 4t + t^2$.

$$
\begin{tikzcd}
& 321 \arrow[dl] \arrow[dr] & \\
231 \arrow[d] & & 312 \arrow[d] \\
132 \arrow[dr] & & 213 \arrow[dl] \\
& 123 &
\end{tikzcd}
$$

1.6 Permutohedron

A **permutahedron** is defined as

$$P = \operatorname{conv}\{(w_1, w_2, \cdots, w_n) \mid w \text{ is a permutation of } [n]\}$$

where we mean a sequence of numbers $1, \cdots, n$ as a permutation instead of an element of the symmetric group S_n in above definition, and $(w_1, \cdots, w_n)$ is a point in $\mathbb{R}^n$, with components the 1-st, 2-nd, $\cdots$, n-th number of the sequence w.

Alternatively, we can say $w \in S_n$ but mean the second row of its matrix notation instead of meaning its cycle notation. For example, The cycle notation $(1\ 2) \in S_3$ is represented in matrix as

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

and is then abbreviated as $w = (2\ 1\ 3) \in S_3$ with $w_1 = 2$, $w_2 = 1$, and $w_3 = 3$.

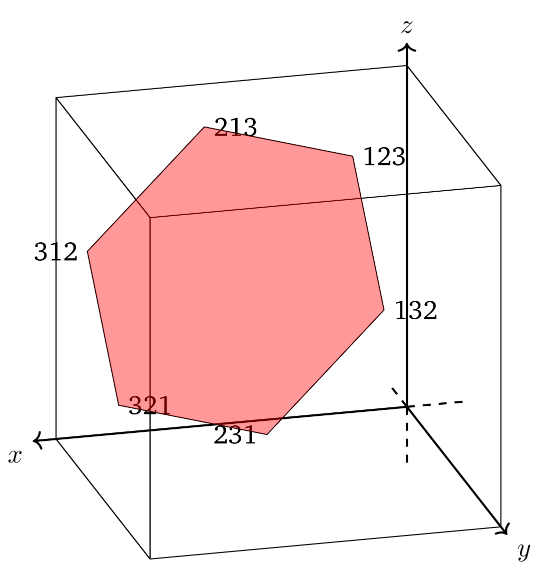


Figure 1.3: conv(S_3)

Just as conv(S_3) can be represented in two dimensions, Figure 1.6 gives the three-dimensional representation of conv(S_4) in four dimensions.

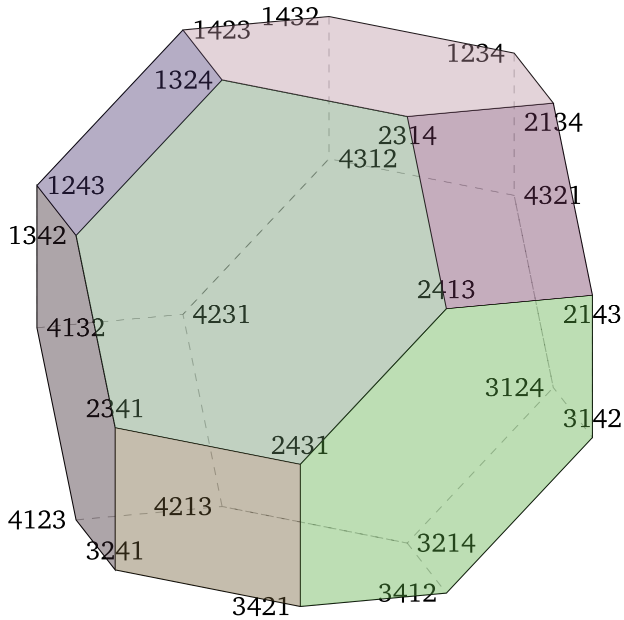


Figure 1.4: conv(S_4)

Recall: There is a method to compute the h-vector of a simple polytope.
(1) Find $\lambda : \mathbb{R}^n \to \mathbb{R}$ linear such that $\lambda(u) \neq \lambda(v)$ for all edges $[u, v]$ of P.
(2) For each vertex u, define $\beta(u) :=| \{v \in \text{verts}(P) \mid [u, v] \text{ is an edge}, \lambda(v) > \lambda(u)\} |$.

Theorem 1.6.1. Let P be a simple polytope. Then $h_P(t) = \sum_{u \in \text{verts}(P)} t^{\beta(u)}$.

Let S_n denote the set of permutations of $[n]$. A **descent** of $w \in S_n$ is an $i \in [n-1]$ such that $w(i) > w(i+1)$. The **Eulerian number** $A(d, i) :=| \{w \in S_n \mid w \text{ has exactly } i \text{ descents } \} |$.

Proposition 1.6.2. Let $\Pi_d := \text{conv}\,(S_n)$ be the permutohedron. The h-polynomial is $h_{\Pi_d}(t) = \sum_{i=0}^{d-1} A(d, i)t^i$.

Proof. Let us choose a linear function $\lambda(x) = \lambda_1 x_1 + \cdots + \lambda_n x_n$ with $\lambda_1 << \cdots << \lambda_n$. Then λ satisfies condition (1). Let $v = (i, i+1)w$ be adjacent to w. Then, $\lambda(v) > \lambda(w) \Leftrightarrow \lambda_{w^{-1}(i)} > \lambda_{w^{-1}(i+1)} \Leftrightarrow w^{-1}(i) > w^{-1}(i+1) \Leftrightarrow i$ is a descent of w^{-1}. Thus, $\beta(w) = \#\left\{i \mid i \text{ is a descent of } w^{-1}\right\}$ and the result follows. □

Exercise 1.6.3. Prove that verts $(\Pi_d) = S_n$ and the edges of Π_d are given by $[w, (i, i+1)w]$ for some i.

1.7 Dual/Polar Polytopes

This is based on [1] section 2.3.

Definition 1.7.1. The polar of $P \subseteq \mathbb{R}^d$ is the set

$$P^{\Delta} = \left\{c \in \mathbb{R}^d \mid \forall x \in P, \langle x, c\rangle \leqslant 1\right\}$$

Notation: given $m \in \mathbb{R}^d$ and $b \in \mathbb{R}$, let $H_{m,b} = \left\{x \in \mathbb{R}^d \mid \langle m, x\rangle \leqslant b\right\}$.

Example 1.7.2. Let $P = \text{conv}\{(0,2),(-1,1),(-1,0),(0,-1),(1,-1)\}$. Then $P^\Delta \subseteq H_{m,1}$ for all $m \in P$. In fact, taking the vertices is enough so that

$$\begin{aligned} P^\Delta &= \left\{(x,y) \in \mathbb{R}^d \mid 2y, -x+y, -x, -y, x-y \leqslant 1\right\} \\ &= \text{conv}\{(-.5,.5),(-1,0),(-1,-1),(0,-1),(1.5,.5)\} \end{aligned}$$

The following figure shows P and P^Δ.

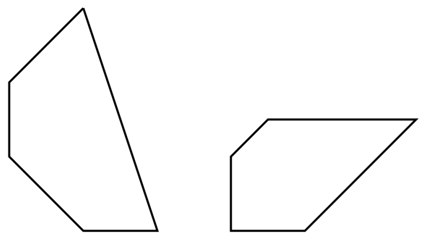


Figure 1.5: P and P^Δ.

Example 1.7.3. Let $P = \text{conv}\{(0,1),(1,1),(1,0)\}$. Then $P^\Delta = \left\{c \in \mathbb{R}^2 \mid c_1 + c_2 \leqslant 1, c_1 \leqslant 1, c_2 \leqslant 1\right\}$. The following figure shows P and P^Δ.

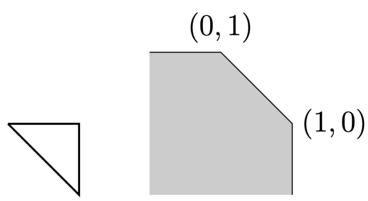


Figure 1.6: P and P^Δ.

Theorem 1.7.4 (ziegler Theorem 2.11)**.**
(i) $P \subseteq Q$ implies $Q^\Delta \subseteq P^\Delta$.
(ii) $P \subseteq P^{\Delta\Delta}$.
(v) If $0 \in P$, then $P = P^{\Delta\Delta}$.
(vi) If $0 \in \text{int}(P)$ and $P = \text{conv}(V)$, then $P^\Delta = \bigcap_{v\in V} H_{v,1}$.
(vii) If $P = \{x \mid Ax \leqslant \mathbb{1}\}$, then P^Δ is the convex hull of the rows of A.

Example 1.7.5. The cube is the polar to the octahedron. To observe (vi) and (vii), note that the vertices of the octahedron are $(\pm 1,0,0),(0,\pm 1,0),(0,0,\pm 1)$. This gives us the inequalities of the cube.

Example 1.7.6. The cube is the polar to the octahedron. To observe (vi) and (vii), note that the vertices of the octahedron are $(\pm 1,0,0),(0,\pm 1,0),(0,0,\pm 1)$. This gives us the inequalities of the cube.

Remark 1.7.7 (For those taking toric variety)**.** A different way to define the polar is as

$$P^* = \left\{c \in \mathbb{R}^d \mid \forall x \in P, \langle x, c\rangle \geqslant -1\right\}.$$

This is prefered in toric geometry. The polytopes P^Δ and P^* are related by the linear isomorphism $\varphi(x) = -x$. This is because P^Δ is the convex hull of the outer normals of the facets of P whereas P^* is the convex hull of the inner normals.

proof of the theorem. (i)-(ii) Exercises.

(v) By (ii) we only need to show that $P^{\Delta\Delta} \subseteq P$. Suppose that $q \in P^{\Delta\Delta}$ but $q \notin P$. Let $c \in P^{\Delta}$. By definition, $\langle c, q\rangle \leqslant 1$. Since $q \notin P$, there exists a hyperplane separating q from P. Suppose that $H_{m,b}$ is such that $q \notin H_{m,b}$ and $P \subseteq H_{m,b}$. Further assume that the boundary of $H_{m,b}$ is disjoint from P. Since $0 \in P$, we have that $b > 0$. It follows that $\langle m/b, x\rangle < b/b = 1$ for all $x \in P$, i.e., $m/b \in P^{\Delta}$. Since $q \in P^{\Delta\Delta}$, we have that $\langle m/b, q\rangle \leqslant 1$. However, this contradicts that $q \notin H_{m,b}$.

(vi) Since $V \subseteq P$, it follows that $P^{\Delta} \subseteq \bigcap_{v \in V} H_{v,1}$. For the opposite containment, let a be such that $\langle a, v\rangle \leqslant 1$ for all $v \in V$. Suppose that $\langle a, x\rangle > 1$ for some $x \in P$. Since the linear functional $\langle a, -\rangle$ is maximized at the face P_a we can take a vertex of P_a which also has to be in V. Now, $\langle a, v\rangle \geqslant \langle a, x\rangle > 1$, contradicting $\langle a, v\rangle \leqslant 1$. □

Next we want to compare the face lattices of P and P^{Δ}.

Example 1.7.8. Let P be the cube and P^{Δ} be the octahedron. We can see that the face lattices are opposites.

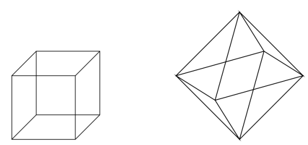


Figure 1.7: P and P^{Δ}.

Definition 1.7.9. Let $(S, \leqslant)$ be a poset. The **opposite poset** $(S, \preceq)$ is defined by $x \preceq y$ if and only if $y \leqslant x$.

Proposition 1.7.10. The face lattice of P^{Δ} is the opposite of the face lattice of P.

This corollary is a consequence of the following theorem.

Theorem 1.7.11 (ziegler, Theorem 2.12)**.** Let $P = \operatorname{conv}(V) = \{x \mid Ax \leqslant \mathbb{1}\}$ and consider a face $F = \operatorname{conv}(V') = \{x \mid A''x \leqslant \mathbb{1}, A'x = \mathbb{1}\}$, where A', A'' together form the rows of A. Then P^{Δ} has a dual face $F^{\Delta} = \operatorname{conv}(\text{ rows of } A') = \{a \mid aV'' \leqslant 1, aV' = 1\}$. Moreover, every face of P^{Δ} is of this form.

We wish to prove:

Theorem 1.7.12 (ziegler, Theorem 2.11)**.** (vii) If $P = \{x \mid Ax \leqslant \mathbb{1}\}$, then P^{Δ} is the convex hull of the rows of A.

First, we need to introduce a version of the Farkas Lemma.

Lemma 1.7.13 (Farkas Lemma III)**.** . Let $A \in \mathbb{R}^{m\times d}, z \in \mathbb{R}^m, a \in \mathbb{R}^d$, and $z_0 \in \mathbb{R}$. The polyhedron $P = \left\{x \in \mathbb{R}^d \mid Ax \leqslant z\right\}$ is nonempty if and only if
(1) there exists a vector $c \geqslant 0$ such that $cA = a$ and $\langle c, z\rangle \leqslant z_0$, or
(2) there exists a vector $c \geqslant 0$ such that $cA = 0$ and $\langle c, z\rangle < 0$, or both.

Proof of (vii). The containment $\supseteq$ is straightforward since every row of A has to be in P^{Δ}. For the opposite containment, let $a \in P^{\Delta}$. Since $P \neq \varnothing$ and condition (2) cannot hold, Farkas Lemma III implies that there exists $c \geqslant 0$ such that $cA = a$ and $\langle c, \mathbb{1}\rangle \leqslant 1$. This is close, but what we really need is $c' \geqslant 0$ such that $c'A = a$ and $\langle c', \mathbb{1}\rangle = 1$. To do so, we will find $d = c' - c \geqslant 0$ with $dA = 0$ and $\langle d, \mathbb{1}\rangle = 1 - c > 0$ which leads us to the desired c' by scaling. Farkas Lemma II says that we can find d unless there exist x, y such that

$$Ax = \mathbb{1}y \quad \text{and} \quad y < 0.$$

However, this would imply that for all $\lambda > 0, \lambda x \in P$ contradicting that P is bounded.

Chapter 2

Graph of Polytopes

2.1 $G(P)$ and linear programming

Let P be a convex polytope. The vertices and the edges of P form an abstract, finite, undirected, simple graph, called the **graph of** P and denoted by $G(P)$.

For every face $F \in L(P)$, we denote by $G(F)$ the **induced subgraph of** $G(P)$ on the subset $\operatorname{vert}(F) \subseteq \operatorname{vert}(P)$ of the vertices of $G(P)$, that is, the graph of all vertices in F, and all edges of P between them. This coincides with the graph of F, if F is itself considered as a polytope.

Definition 2.1.1. A linear function $\lambda : \mathbb{R}^d \to \mathbb{R}$ is in **general position** with respect to a polytope P if for all $u, v \in$ Vert(P), $\lambda(u) \neq \lambda(v)$.

Definition 2.1.2. We will consider **orientations** of $G(P)$, which assign a direction to every edge. An orientation is **acyclic** if there is no directed cycle in it. This implies (because all our graphs are finite) that there is a sink: a vertex that does not have an edge directed away from it. (Proof: Start at any vertex, and keep on walking along directed edges until you close a directed cycle or get stuck in a sink.)

If λ is in general position with repsect to P, then it induces an orientation of $G(P)$. Concretely, $u \to v$ if $\lambda(u) > \lambda(v)$.

Proposition 2.1.3. The orientation of $G(P)$ induced by λ in general position is acyclic and has a unique sink. Moreover, λ is maximized over P at the sink.

Proof. If there was a cycle $v_1 \to \cdots \to v_k \to v_1$, then $\lambda\left(v_1\right) > \cdots > \lambda\left(v_k\right) > \lambda\left(v_1\right)$, a contradiction. Every acyclic graph has a sink. Suppose t is a sink. Let $N(t) = \{v \in \operatorname{verts}(P) \mid [t, v]$ edge $\}$ be the neighbors of t. Recall that the vertex figure P/t is obtained by cutting P by a hyperplane that separates v from the other vertices of P. The vertices of P/t are in 1-1 correspondence with the elements of $N(t)$. It follows that $P \subseteq t + \operatorname{cone}(v - t \mid v \in N(t))$. Now, since $\lambda(v) \leqslant \lambda(t)$ for all $v \in N(t)$ we have that given $p \in P$,

$$\lambda(p) = \lambda(t) + \sum_{v \in N(t)} \lambda(v - t) \leqslant \lambda(t),$$

since $\lambda(v - t) < 0$ for all $v \in N(t)$. It follows that t is the unique sink and that it maximizes λ. □

This proposition gives us a method to maximize λ over P.

Proposition 2.1.4 (Dantzig's simplex algorithm)**.**
(1) Start at a vertex v.

(2) If v is a sink, then stop.
(3) Otherwise, move to a neighbor w of v such that $\lambda(w) > \lambda(v)$.

.

This hides the fact that finding a vertex of P is nontrivial. The full simplex algorithm takes care of this.

2.2 The diameter of a polytope.

Definition 2.2.1. The **diameter** $\delta(P)$ of P is the the smallest δ such that any two vertices in P can be connected by a path with at most δ edges.

The diameter of a polytope is a measure of how hard it is to optimize a linear function over it using the simplex algorithm. Concretely, it gives a lower bound on the number of iterations needed.

Let $\Delta(d,n)$ be the maximum diameter of a d-dimensional polytope with at most n facets.

Example 2.2.2. $\Delta(2,n) = \lfloor n/2 \rfloor$.

Question: What is the behavior of $\Delta(d,n)$?

Conjecture(Hirsch Conjecture, '57). $\Delta(d,n) \leqslant n-d$.

This was disproven:

Theorem 2.2.3 (Santos, '10)**.** The Hirsch Conjecture is false. Namely, there exists a 43-dimensional polytope with 86 facets such that $\delta(P) \geqslant 44$.

Conjecture(Polynomial Hirsch Conjecture). There is a polynomial $f(n)$ such that the diameter of every polytope with n facets is bounded above by $f(n)$.

2.3 Simple polytope and graph

Theorem 2.3.1 (Perles '70 - conjecture, Blind-Mani '87)**.** If P is simple then $G(P)$ determines P up to combinatorial equivalence.

In other words, if two simple polytopes have isomorphic graphs, then their face lattices are isomorphic as well.

We will discuss Kalai's simple proof of this result. A key observation will be that if P is simple, then for any vertex v any k edges incident to v determine a face.

proof of the theorem. Let $\mathcal{O}(P)$ be the set of all orientations of $G(P)$. Let us say that $\mathcal{O} \in \mathcal{O}(P)$ is **good** if for all faces F of P, $\mathcal{O}|_{\mathrm{verts}(F)}$ has a unique sink. (Otherwise, we say that $\mathcal{O}$ is **bad**.) We say that a graph is k**-regular** if all its vertices have degree k. The result follows from the following two claims.

Claim 1. Let $U \subseteq \mathrm{verts}(P)$. Then U is the vertex set of a k-dimensional face of P if and only if $G(P)|_U$ is k-regular and there exists a good orientation $\mathcal{O}$ for which U is a downset. Here the last condition means that if $v \in U$ and $u \to v$ is an edge of $\mathcal{O}$, then $u \in U$.

Let $f^{\mathcal{O}} = \sum_{v \in \mathrm{verts}(P)} 2^{\text{indeg(v)}}$.

Claim 2:

$$\{\text{ good orientations }\} = \left\{\text{ orientations with } \max f^{\mathcal{O}}\right\}.$$

Given $\mathcal{O} \in \mathcal{O}(P)$, note that

$$
\begin{aligned}
\mid \{\text{ nonempty faces of } P\} \mid \leqslant & \mid \{(F, v) \mid F \text{ face of } P \text{ and } v \text{ is a sink vertex of } F\} \mid \\
& = \sum_{v \in \operatorname{verts}(P)} \mid \{F \mid v \text{ is a sink vertex of } F\} \mid \\
& = \sum_{v \in \operatorname{verts}(P)} 2^{\operatorname{indeg}(v)}
\end{aligned}
$$

where the last equality follows from the fact that P is simple. More concretely, for any vertex v any k edges incident to v determine a face. It follows that if v has indegree k, then there are 2^k faces of P with v as a sink. Moreover, we have equality if and only if $\mathcal{O}$ is good. This gives us the most recent claim.

We are now ready to prove the first claim.

$\Rightarrow$ Suppose that U are the vertices of a face F. Since P is simple, it is immediate that $\left.G(P)\right|_U$ is k-regular. Let λ be such that F minimizes λ over P. This λ may not be in general position, but we can perturb by a small amount to be so. The resulting orientation is good and U is a *downset*.

$\Leftarrow$ Suppose that U and $\mathcal{O}$ are as desired. Since $\mathcal{O}_U$ is acyclic, let x be a sink and note that it has indegree k. Let F be the k-dimensional face determined by these k edges. Since $\mathcal{O}$ is good, x is the unique sink of F. Since $u \rightarrow x$ for all vertices along these k edges and U is a downset, then $\operatorname{verts}(F) \subseteq U$. Now, we must have that verts $(F)=U$ since both $\left.G\right|_U$ and $\left.G\right|_{\operatorname{verts}(F)}$ are connected and k-regular. $\square$

Theorem 2.3.2 (Balinksi's Theorem)**.** A graph is connected if there is a path between any two vertices.

Definition 2.3.3. A graph is d**-connected** if it stays connected after removing any $\leqslant d-1$ vertices (and their incident edges).

Theorem 2.3.4. If P is d-dimensional, then $G(P)$ is d-connected.

In particular, this says that the degree of any vertex is $\geqslant d$.

Proof. To ease the proof, suppose that $P \subseteq \mathbb{R}^d$. Let $V=\operatorname{verts}(P)$ and $S \subseteq V$ be such that $|S| \leqslant d-1$. To show that $G(P)-S$ is connected we use induction on d. The base case $d=1$ is immediate. For the inductive case, let $L=\operatorname{span}(S)$ and consider two cases.

(1) Suppose L doesn't intersect the interior of P. Then S are the vertices of a face $P_c \subsetneq P$. Consider the face P_{-c} and the orientation of $G(P)$ given by the function $\lambda(x)=\langle-c, x\rangle$. By the argument we used in the proof of Proposition 2.1.3, we have have that every vertex is either in P_{-c},or it has a neighbor $x \notin S$ whose $\langle c, x\rangle$-value is smaller. Thus, there is a c decreasing path from any vertex in $V \backslash S$ to a vertex in P_{-c}. By induction, $G\left(P_{-c}\right)$ is connected and we are done.

(2) Suppose L intersects the interior of P. Let $H=\left\{x \mid\langle c, x\rangle=c_0\right\}$ be a hyperplane containing S and at least one $v \in V \backslash S$. Note that since $L \subseteq H$, then H also intersects the interior of P. This is possible because every set of d points is contained in a hyperplane Consider the faces P_c and P_{-c}. Let $P^{+}=\left\{x \in P \mid\langle c, x\rangle \geqslant c_0\right\}$ and $P^{-}=\left\{x \in P \mid\langle c, x\rangle \leqslant c_0\right\}$. Note that every vertex of P^{+}has a c-increasing path to P_c. Since $G\left(P_c\right)$ is connected, by induction, it follows that $G\left(P^{+}\right) \backslash S$ is connected. Similarly, $G\left(P^{-}\right) \backslash S$ is connected. Since v is in both graphs, we conclude that $G(P) \backslash S$ is connected. $\square$

Question: Can we characterize the graphs of polytopes?

Theorem 2.3.5 (Steinitz' Theorem)**.** G is the graph of a 3-dimensional polytope if and only if it is simple, planar, and 3-connected (Simple means that it has no loops or multiple edges.)

Proof. Proof of $\Rightarrow$. Let G be the graph of a 3-dimensional polytope P. It is immediate that it is simple. Also, Balinksi's Theorem implies it is 3-connected. Last, it is planar by blowing up a facet.

Proof of $\Leftarrow$. See [2, Chapter 4]. $\square$

Remark 2.3.6. No similar theorem is known, and it seems that no similarly effective theorem is possible, in higher dimensions.

Chapter 3

The Ehrhart Theory

We shift to [4] for the main reference. The main theme is to count the number of integer points inside a polytope. We begin with some examples.

A convex polytope $\mathcal{P}$ is called **integral** if all of its vertices have integer coordinates, and $\mathcal{P}$ is called **rational** if all of its vertices have rational coordinates. A **unit d-cube**

$$\begin{aligned} \square &:= [0,1]^d = \{(x_1, x_2, \cdots, x_d) \in \mathbb{R}^d : \text{ all } x_k = 0 \text{ or } 1\} \\ &= \left\{(x_1, x_2, \ldots, x_d) \in \mathbb{R}^d : 0 \leqslant x_k \leqslant 1 \text{ for all } k = 1, 2, \ldots, d\right\} \end{aligned}$$

We now compute the discrete volume of an integer dilate of $\square$. That is, we seek the number of integer points $t\,\square \cap \mathbb{Z}^d$ for all $t \in \mathbb{Z}_{>0}$. Here $t\mathcal{P}$ denotes the dilated polytope $\{(tx_1, tx_2, \ldots, tx_d) : (x_1, x_2, \ldots, x_d) \in \mathcal{P}\}$ for a polytope $\mathcal{P}$. What is the discrete volume of $\mathcal{P} = \square$? We dilate by the positive integer t, as depicted in Figure 3, and count:

$$L_{\mathcal{P}}(t) := \#\left(t\mathcal{P} \cap \mathbb{Z}^d\right) = \#\left(t\,\square \cap \mathbb{Z}^d\right) = \#\left([0,t]^d \cap \mathbb{Z}^d\right) = (t+1)^d,$$

a polynomial in the integer variable t. Notice that the coefficients of this polynomial are the binomial coefficients. The number of interior integer points in $t\square^\circ$ is $L_{\square^\circ}(t) = \#\left(t\,\square^\circ \cap \mathbb{Z}^d\right) = \#\left((0,t)^d \cap \mathbb{Z}^d\right) = (t-1)^d$. Notice that this polynomial equals $(-1)^d L_\square(-t)$, the evaluation of the polynomial $L_\square(t)$ at negative integers, up to a sign.

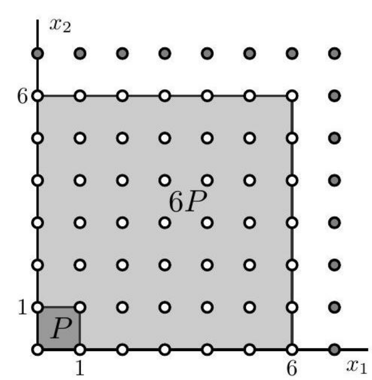


Figure 3.1: The 6^{th} dilate of $\square$ in dimension 2 .

The generating function of $L_{\mathcal{P}}$ is called **Ehrhart series** of $\mathcal{P}$:

$$\mathrm{Ehr}_{\mathcal{P}}(z) := 1 + \sum_{t\geqslant 1} L_{\mathcal{P}}(t) z^t$$

The Ehrhart series of $\mathcal{P} = \square$ takes on a special form.

$$\begin{aligned}\mathrm{Ehr}_{\square}(z) = 1 + \sum_{t\geqslant 1}(t+1)^d z^t = \sum_{t\geqslant 0}(t+1)^d z^t = \frac{1}{z}\sum_{t\geqslant 1} t^d z^t \\ = \frac{\sum_{k=1}^{d} A(d,k) z^{k-1}}{(1-z)^{d+1}}.\end{aligned}$$

where $A(d,k)$ is the **Eulerian number**, which counts the number of permutations with exactly k descents.

The standard simplex Δ in dimension d is

$$\operatorname{conv}\{\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d\} = \left\{(x_1, x_2 \ldots, x_d) \in \mathbb{R}^d : x_1 + x_2 + \cdots + x_d \leqslant 1 \text{ and all } x_k \geqslant 0\right\}$$

(we now include zero as a vertex now). The dilate $t\Delta$ is

$$t\Delta = \left\{(x_1, x_2, \ldots, x_d) \in \mathbb{R}^d : x_1 + x_2 + \cdots + x_d \leqslant t \text{ and all } x_k \geqslant 0\right\}$$

The lattice-point enumerator of Δ is the polynomial $L_\Delta(t) = \left(\begin{array}{c} d+t \\ d \end{array}\right)$. Its evaluation at negative integers yields $(-1)^d L_\Delta(-t) = L_{\Delta^\circ}(t)$. The Ehrhart series of Δ is $\mathrm{Ehr}_\Delta(z) = \frac{1}{(1-z)^{d+1}}$.

3.1 Triangulations

Because most of the proofs that follow work like a charm for a simplex, we first dissect a polytope into simplices. This dissection is captured by the following definition.

A triangulation of a convex d-polytope $\mathcal{P}$ is a finite collection T of d-simplices with the following properties:

- $\mathcal{P} = \bigcup_{\Delta \in T} \Delta$.
- For every $\Delta_1, \Delta_2 \in T$, $\Delta_1 \cap \Delta_2$ is a face common to both Δ_1 and Δ_2.

Figure 3.1 exhibits two triangulations of the 3-cube. We say that $\mathcal{P}$ can be triangulated using no new vertices if there exists a triangulation T such that the vertices of every $\Delta \in T$ are vertices of $\mathcal{P}$.

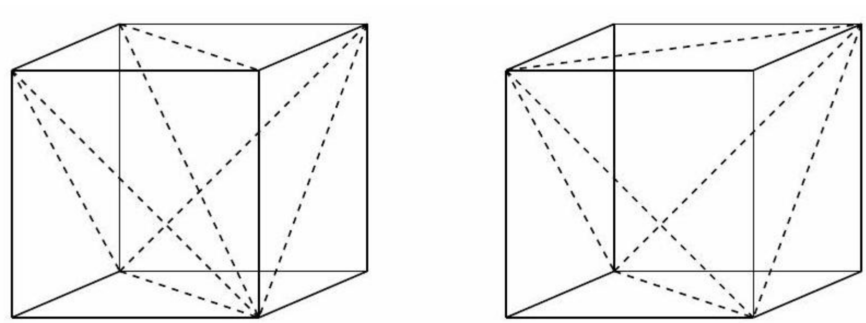


Figure 3.2: Two (very different) triangulations of the 3-cube.

Theorem 3.1.1 (Existence of triangulations). Every convex polytope can be triangulated using no new vertices.

This theorem seems intuitively obvious, but it is not entirely trivial to prove.

In order to prove this, let us consider regular subdivisions.

Let $P = \operatorname{conv}(V) \subseteq \mathbb{R}^d$. Choose $h : V \to \mathbb{R}$ and let $P' = \operatorname{conv}\left\{ \left[\begin{array}{c} v \\ h(v) \end{array} \right] \middle| \, v \in V \right\}$. We say that a face F of P' is lower if $F = P'_c$ for some $c \in \mathbb{R}^{d+1}$ with $c_{d+1} < 0$.

Let $\pi : \mathbb{R}^{d+1} \to \mathbb{R}^d$ be the projection onto the first d coordinates.

Proposition 3.1.2. The set $\{\pi(F) \mid F \text{ lower face of } P'\}$ is a subdivision of P. If h is generic, then this subdivision is a triangulation.

Proof. We will only provethe second claim. Suppose $P \subseteq \mathbb{R}^d$ is d-dimensional. First, we will show that each lower facet of P' is a simplex, i.e., the convex hull of $d+1$ affinely independent vectors. Suppose we have $d+1$ affinely independent vertices of $P, v_1, \ldots, v_{d+1}$. Let $H \subseteq \mathbb{R}^{d+1}$ be the hyperplane given by the equation

$$\left[\begin{array}{ccccc} 1 & 1 & \cdots & 1 & 1 \\ v_1 & v_2 & \cdots & v_{d+1} & \boldsymbol{x} \\ h(v_1) & h(v_2) & \cdots & h(v_{d+1}) & x_{d+1} \end{array} \right]$$

Note that for all i, $\left[\begin{array}{c} v_i \\ h(v_i) \end{array} \right] \in H$. Also, note that if we fix $\boldsymbol{x}$, then there is a unique x_{d+1} that makes this equation hold. Thus, if $v \neq v_i$ for all i, then since we chose the $h(v)$ at random, then $\left[\begin{array}{c} v \\ h(v) \end{array} \right]$ is not in H. This proves that P' is simplicial, so all of its faces are simplices.

Next, we will show that $\bigcup_{\Delta \in T} \Delta = P$. It suffices to show that $\operatorname{int}(P) \subseteq \bigcup_{\Delta \in T} \Delta$. Let $x \in \operatorname{int}(P)$ and consider the vertical line $\mathcal{L} \subseteq \mathbb{R}^{d+1}$ through x. Since $\mathcal{L} \cap \operatorname{int}(P') \neq \varnothing$, then $\mathcal{L} \cap P'$ is a line segment with endpoints (x, y) and $(x, z), y < z$. Since $(x, y) \in \partial P'$, then $(x, y) \in P'_{(c, c_{d+1})}$ for some (c, c_{d+1}). Note then that $\langle (c, c_{d+1}), (x, y) \rangle > \langle (c, c_{d+1}), (x, z) \rangle$ and since $y < z$ this can only hold if $c_{d+1} < 0$. It follows that P'_c is a lower face and $x \in \pi(P'_c)$.

The last property is left as an exercise

□

3.2 The Ehrhart Series of an Rational Polytope

By now, we have computed several instances of counting functions by setting up a generating function that fits the particular problem in which we are interested. In this subsection, we set up such a generating function for the latticepoint enumerator of an arbitrary rational polytope. Such a polytope is given by its hyperplane description as an intersection of half-spaces and hyperplanes. The half-spaces are algebraically given by linear inequalities, the hyperplanes by linear equations. If the polytope is rational, we can choose the coefficients of these inequalities and equations to be integers (Exercise). To unify both descriptions, we can introduce slack variables to turn the half-space inequalities into equalities. Furthermore, by translating our polytope into the nonnegative orthant (we can always shift a polytope by an integer vector without changing the lattice-point count), we may assume that all points in the polytope have nonnegative coordinates. In summary, after a harmless integer translation, we can describe every rational polytope $\mathcal{P}$ as

$$\mathcal{P} = \left\{ \mathbf{x} \in \mathbb{R}^d_{\geqslant 0} : \mathbf{A}\mathbf{x} = \mathbf{b} \right\}$$

for some integral matrix $\mathbf{A} \in \mathbb{Z}^{m \times d}$ and some integer vector $\mathbf{b} \in \mathbb{Z}^m$. (Note that d is not necessarily the dimension of $\mathcal{P}$.) To describe the t^{th} dilate of $\mathcal{P}$, we simply scale a point $\mathbf{x} \in \mathcal{P}$ by $\frac{1}{t}$, or alternatively, multiply $\mathbf{b}$ by t :

$$t\mathcal{P} = \left\{\mathbf{x} \in \mathbb{R}_{\geqslant 0}^d : \mathbf{A}\frac{\mathbf{x}}{t} = \mathbf{b}\right\} = \left\{\mathbf{x} \in \mathbb{R}_{\geqslant 0}^d : \mathbf{A}\mathbf{x} = t\mathbf{b}\right\}$$

Hence the lattice-point enumerator of $\mathcal{P}$ is the counting function

$$L_{\mathcal{P}}(t) = \#\left\{\mathbf{x} \in \mathbb{Z}_{\geqslant 0}^d : \mathbf{A}\mathbf{x} = t\mathbf{b}\right\}$$

Consider the polytope $\mathcal{P}$ given by (3.2), we denote the columns of $\mathbf{A}$ by $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_d$. Let $\mathbf{z} = (z_1, z_2, \ldots, z_m)$ and expand the function

$$\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)\mathbf{z}^{t\mathbf{b}}}$$

in terms of geometric series:

$$\left(\sum_{n_1 \geqslant 0} \mathbf{z}^{n_1\mathbf{c}_1}\right)\left(\sum_{n_2 \geqslant 0} \mathbf{z}^{n_2\mathbf{c}_2}\right)\cdots\left(\sum_{n_d \geqslant 0} \mathbf{z}^{n_d\mathbf{c}_d}\right)\frac{1}{\mathbf{z}^{t\mathbf{b}}}.$$

Here we use the abbreviating notation $\mathbf{z}^{\mathbf{a}} := z_1^{a_1} z_2^{a_2} \cdots z_m^{a_m}$ for the vectors $\mathbf{z} = (z_1, z_2, \ldots, z_m) \in \mathbb{C}^m$ and $\mathbf{a} = (a_1, a_2, \ldots, a_m) \in \mathbb{Z}^m$. In multiplying out everything, a typical exponent will look like

$$n_1\mathbf{c}_1 + n_2\mathbf{c}_2 + \cdots + n_d\mathbf{c}_d - t\mathbf{b} = \mathbf{A}\mathbf{n} - t\mathbf{b}$$

where $\mathbf{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{Z}_{\geqslant 0}^d$. That is, if we take the constant term of our generating function (3.2), we are counting integer vectors $\mathbf{n} \in \mathbb{Z}_{\geqslant 0}^d$ satisfying

$$\mathbf{A}\mathbf{n} - t\mathbf{b} = \mathbf{0}, \quad \text{that is,} \quad \mathbf{A}\mathbf{n} = t\mathbf{b}.$$

So this constant term will pick up exactly the number of lattice points $\mathbf{n} \in \mathbb{Z}_{\geqslant 0}^d$ in $t\mathcal{P}$:

Theorem 3.2.1 (Euler's generating function)**.** Suppose the rational polytope $\mathcal{P}$ is given by (3.2). Then the lattice-point enumerator of $\mathcal{P}$ can be computed as follows:

$$L_{\mathcal{P}}(t) = \operatorname{const}\left(\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)\mathbf{z}^{t\mathbf{b}}}\right)$$

We finish this section with rephrasing this constant-term identity in terms of Ehrhart series.

Corollary 3.2.2. Suppose the rational polytope $\mathcal{P}$ is given by (3.2). Then the Ehrhart series of $\mathcal{P}$ can be computed as

$$\operatorname{Ehr}_{\mathcal{P}}(x) = \operatorname{const}\left(\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)\left(1-\frac{x}{\mathbf{z}^{\mathbf{b}}}\right)}\right)$$

Proof. By above theorem,

$$\begin{aligned}
\operatorname{Ehr}_{\mathcal{P}}(x) &= \sum_{t \geqslant 0} \operatorname{const}\left(\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)\mathbf{z}^{t\mathbf{b}}}\right)x^t \\
&= \operatorname{const}\left(\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)}\sum_{t \geqslant 0}\frac{x^t}{\mathbf{z}^{t\mathbf{b}}}\right) \\
&= \operatorname{const}\left(\frac{1}{\left(1-\mathbf{z}^{\mathbf{c}_1}\right)\left(1-\mathbf{z}^{\mathbf{c}_2}\right)\cdots\left(1-\mathbf{z}^{\mathbf{c}_d}\right)}\frac{1}{1-\frac{x}{\mathbf{z}^{\mathbf{b}}}}\right).
\end{aligned}$$

□

3.3 Ehrhart Theory of Integral Polytopes

We wish to understand L_P and Ehr_P for polytopes P such that $\mathrm{verts}(P) \subseteq \mathbb{Z}^d$.

Ehrhart theory for cones

Given $S \subseteq \mathbb{R}^d$, let

$$\sigma_S(\mathbf{z}) = \sum_{m \in S \cap \mathbb{Z}^d} \mathbf{z}^m.$$

Note that $\sigma_S(1) = L_S(t) = \left|S \cap \mathbb{Z}^d\right|$.

Chapter 4

Fourier Analysis on Polytopes

Long presentation materials:

The **Poisson summation formula** tells us that for any "sufficiently nice" function $f : \mathbb{R}^d \to \mathbb{C}$ we have:

$$\sum_{n\in\mathbb{Z}^d} f(n) = \sum_{\xi\in\mathbb{Z}^d} \hat{f}(\xi).$$

In particular, if we were to naively set $f(n) := 1_{\mathcal{P}}(n)$, the indicator function of a polytope $\mathcal{P}$, then we would get:

$$\sum_{n\in\mathbb{Z}^d} 1_{\mathcal{P}}(n) = \sum_{\xi\in\mathbb{Z}^d} \hat{1}_{\mathcal{P}}(\xi), \tag{4.1}$$

which is technically false for functions, due to the fact that the indicator function $1_{\mathcal{P}}$ is discontinuous on $\mathbb{R}^d$. But when we do counting, Knuth says we sometimes don't need to take care of those requirements to use some formulae because they serve as guessing tools. The end justifies the means.

The combinatorial geometric quantity $\left|\mathcal{P} \cap \mathbb{Z}^d\right|$ may be regarded as a **discrete volume** for $\mathcal{P}$. From the definition of the indicator function of a polytope, the left-hand-side of (1.4) counts the number of integer points in $\mathcal{P}$, namely we have by definition

$$\sum_{n\in\mathbb{Z}^d} 1_{\mathcal{P}}(n) = \left|\mathcal{P} \cap \mathbb{Z}^d\right|$$

On the other hand, the right-hand-side of (4.1) allows us to compute this discrete volume of $\mathcal{P}$ in a new way. This is great, because it opens a wonderful window of computation for us in the following sense:

$$\left|\mathcal{P} \cap \mathbb{Z}^d\right| = \sum_{\xi\in\mathbb{Z}^d} \hat{1}_{\mathcal{P}}(\xi)$$

We notice that for the $\xi = 0$ term, we have

$$\hat{1}_{\mathcal{P}}(0) := \int_{\mathbb{R}^d} 1_{\mathcal{P}}(x) e^{-2\pi i\langle 0,x\rangle} dx = \int_{\mathcal{P}} dx = \mathrm{vol}(\mathcal{P}),$$

and therefore the discrepancy between the continuous volume of $\mathcal{P}$ and the discrete volume of $\mathcal{P}$ is

$$\left|\mathcal{P} \cap \mathbb{Z}^d\right| - \mathrm{vol}(\mathcal{P}) = \sum_{\xi\in\mathbb{Z}^d-\{0\}} \hat{1}_{\mathcal{P}}(\xi),$$

showing us very quickly that indeed $\left|\mathcal{P} \cap \mathbb{Z}^d\right|$ is a discrete approximation to the classical Lebesgue volume $\operatorname{vol}(\mathcal{P})$, and pointing us to the task of finding ways to evaluate the transform $\hat{1}_P(\xi)$. From the trivial but often very useful identity

$$\hat{1}_{\mathcal{P}}(0) = \operatorname{vol}(\mathcal{P}),$$

Keys: Brion theorems and Euler-Maclaurin Summation

[3] chapter 8, 10

[4] chapter 5, 11, 12 Euler-Maclaurin Summation as the end.

Bibliography

[1] Ziegler, Günter M. *Lectures on Polytopes*, Vol. 152, Springer Science & Business Media, 2012.

[2] Beck, Matthias and Robins, Sinai. *Computing the Continuous Discretely*, Vol. 61, Springer, 2007.

[3] Robins, Sinai. *A Friendly Introduction to Fourier Analysis on Polytopes*, arXiv preprint arXiv:2104.06407, 2021.

[4] Beck, Matthias and Robins, Sinai. *Computing the Continuous Discretely*, Springer, vol. 61, 2007.