Chapter 1

POLYHEDRAL FACTS FOR INTEGER OPTIMIZATION

The purpose of these lecture notes is to provide beginning graduate students and advance undergraduates with a *very quick* introduction to the polyhedral theory needed in integer linear optimization while keeping things fairly formal, i.e. proving (almost) all the claims. These results will be freely use in the integer optimization theory. For our purposes Schrijver and Ziegler are good sources.

Our story takes place inside Euclidean d-dimensional space $\mathbb{R}^d$. We use the traditional Euclidean distance between two points. Given two points x, y. We

will use the common fact that $\mathbb{R}^d$ is a real vector space and thus we know how to add or scale its points. The set of points $[x, y] := \{\alpha x + (1-\alpha)y : 0 \leq \alpha \leq 1\}$ is called *the interval* between x and y. The points x and y are the endpoints of

the interval. We say that set $A \subset \mathbb{R}^d$ is *convex* if for any two points $x, y \in A$ we have that the interval $[x, y]$ is also contained in A. We will assume that the empty set is also convex. Observe that the intersection of convex sets is convex too.

Let $A \subset \mathbb{R}^d$, the *convex hull* of A, denoted by $conv(A)$, is the intersection of all the convex sets containing A. In other words, A is the smallest convex set containing A. The reader can check that the image of a convex set under a linear transformation is again a convex set. An important definition

Definition 1.0.1 *A* polytope *is the convex hull of a finite set of points in* $\mathbb{R}^d$.

Lemma 1.0.2 *For a finite set of points in $\mathbb{R}^d$ $A := \{a_1, a_2, \ldots, a_n\}$ we have that $conv(A)$ equals*

$$\{\sum_{i=1}^{n} \gamma_i a_i : \gamma_i \geq 0 \text{ and } \gamma_1 + \ldots \gamma_n = 1\}$$

Proof: There are two containments that need to be verified: Denote by B the set of finite linear combinations presented above. Clearly A is a subset of B and it is easy to verify is convex, hence $conv(A) \subset B$. Now we must see that $conv(A)$ contains B. We can prove that an element $x = \sum_{i=1}^{n} \gamma_i a_i : \gamma_i \in B$ is in $conv(A)$ by induction on the number of non-zero elements r in the sum. If only $r = 1$ then clearly $x \in A \subset conv(A)$. Suppose there are $r+1$ non-zero coefficients γ's, say $x = \gamma_{j_1} a_{j_1} + \gamma_{j_2} a_{j_2} + \ldots + \gamma_{j_{r+1}} a_{j_{r+1}}$ we can construct $\beta_{j_i} = \gamma_{j_i} / (\gamma_{j_1} + \gamma_{j_2} + \ldots + \gamma_{j_r}$. The β's are all positive and add to one. Using induction we conclude that $y = \beta_{j_1} a_{j_1} + \beta_{j_2} a_{j_2} + \ldots + \beta_{j_{r+1}} a_{j_{r+1}}$ is indeed an element of $conv(A)$. Finally recall that $1 - \gamma_{j_{r+1}}$ equals $\gamma_{j_1} + \gamma_{j_2} + \ldots + \gamma_{j_r}$ hence clearly.

$$x = (1 - \gamma_{j_{r+1}}) y + \gamma_{j_{r+1}} a_{j_{r+1}}$$

Since y and a_{r+1} are in $conv(A)$ and $conv(A)$ is a convex set we conclude that $x \in conv(A)$.

Now it is easier to speak about examples of polytopes!!

1. **Standard Simplex** Let $e_1, e_2, \ldots, e_{d+1}$ be the standard unit vectors in $\mathbb{R}^{d+1}$. The *standard* d-dimensional simplex Δ_d is $conv(\{e_1, \ldots, e_{d+1}\})$. From the above lemma we see that the set is precisely

   $\Delta_d = \{x = (x_1, \ldots, x_{d+1}) : x_i \geq 0 \text{ and } x_1 + x_2 + \ldots + x_{d+1} = 1\}$.

   Note that for a polytope $P = conv(\{a_1, \ldots, a_m\})$ we can define a linear map $f : \Delta_{m-1} \to P$ by the formula $f(\lambda_1, \ldots, \lambda_m) = \lambda_1 a_1 + \ldots + \lambda_m a_m$. Lemma 1.0.2 implies that $f(\Delta_{m-1}) = P$. Hence, every polytope is the image of the standard simplex under a linear transformation. A lot of the properties of the standard simplex are then shared by all polytopes.

2. **Standard Cube** Let $\{u_i : i \in I\}$ be the set of all 2^d vectors in $\mathbb{R}^d$ whose coordinates are either 1 or -1. The polytope $I_d = conv(\{u_i : i \in I\}$ is called the *standard* d-dimensional cube. The images of a cube under linear transformations receive the name of *zonotopes*.

   Clearly $I^d = \{x = (x_1, \ldots, x_d) : -1 \leq x_i \leq 1\}$

3. **Standard Crosspolytope** This is the convex hull of the $2d$ vectors $e_1, -e_1, e_2, -e_2, \ldots, e_d, -e_d$. The 3-dimensional crosspolytope is simply an octahedron.

Recall from linear algebra that a linear function $f : \mathbb{R}^d \to \mathbb{R}$ is given by a vector $c \in \mathbb{R}^d, c \neq 0$. For a number $\alpha \in \mathbb{R}$ we say that $H_\alpha = \{x \in \mathbb{R}^d : f(x) = \alpha\}$ is an *affine hyperplane* or *hyperplane* for short. Note that a hyperplane divides $\mathbb{R}^d$ into two halfspaces $H_\alpha^+ = \{x \in \mathbb{R}^d : f(x) \geq \alpha\}$ and $H_\alpha^- = \{x \in \mathbb{R}^d : f(x) \leq \alpha\}$. Halfspaces are convex sets. So it makes sense to make the following definition:

Definition 1.0.3 *A* polyhedron *in $\mathbb{R}^d$ is the intersection of finitely many half-spaces. In its general form a polyhedron is then a set of the type*

$$P = \{x \in \mathbb{R}^d :< c_i, x > \leq \beta_i\}$$

for some non-zero vectors c_i in $\mathbb{R}^d$ and some real numbers β_i.

Note that the definition includes the existence of equalities. Note also that polyhedra are indeed convex sets. What is the relation between polytopes and polyhedra (yes the plural of polyhedron has an a at the end)? Clearly standard cubes,simplices and crosspolytopes are also polyhedra, but is this the case in general? What one expects is really true:

Theorem 1.0.4 (Weyl-Minkowski theorem) *Every polytope is a polyhedron. Every bounded polyhedron is a polytope.*

This theorem is very important. Having this double way of representing a polytope allows you to prove using one of the representations that would be hard to prove using the alternative representation. For example:

- Every intersection of a polytope with an affine subspace is a polytope
- The intersection of finitely many polytopes is a polytope.
- Every linear projection of a bounded polyhedron is a bounded polyhedron.

In addition, the Weyl-Minkowski theorem is very useful in applications! its existence is key in combinatorial optimization.

1.0.1 Weyl-Minkowski and Polarity

Before we discuss a proof of Weyl-Minkowski theorem we need to introduce a useful operation. To every subset of Euclidean space we wish to associate a convex set. Given a subset A of $\mathbb{R}^d$ the *polar* of A is the set A^o in $\mathbb{R}^d$ defined as the linear functionals whose value on A is not greater than 1, in other words:

$$A^o = \{x \in \mathbb{R}^d :< x, a > \leq 1 \; for \; every \; a \in A\}$$

Another way of thinking of the polar is as the intersection of the halfspaces, one for each element $a \in A$, of the form

$$\{x \in \mathbb{R}^d :< x, a >\leq 1\}$$

Here are two little useful examples: Take L a line in $\mathbb{R}^2$ passing through the origin, what is L^0? Well the answer is the perpendicular line that passes through the origin. If the line does not pass through the origin the answer is different. What is it? Answer: it is a clipped line orthogonal to the given line that passes through the origin. To see without loss of generality rotate the line until it is of the form $x = c$ (because the calculation of the polar boils down to checking angles and lengths between vectors we must get the same answer up to rotation).

What happens with a circle of radius one with center at the origin? Its polar set is the disk of radius one with center at the origin. Next take $B(0, r)$. What is $B(0, r)^o$?

The concept of polar is rather useful. We use the following lemma:

Lemma 1.0.5 *1. If P is a polytope and $0 \in P$, then $(P^o)^o = P$.*

2. Let $P \subset \mathbb{R}^d$ be a polytope. Then P^o is a polyhedron.

Proof:

1. Note $P \subset (P^o)^o$ (immediate from the definition of polar). We just need to check that $(P^o)^o \subset P$. Suppose $y \in (P^o)^o$ but $y \notin P$. Consider a hyperplane that separates P and y, namely $< x, c >= \alpha$ that such that $< y, c >< \alpha$ and $< p, c > \alpha$ for $p \in P$ (such hyperplane exist and it is a very deep fact you saw in calculus). Since $0 \in P$ we have that $\alpha > 0$. Therefore letting $b = \alpha^-1c$ we get $< b, x >< 1$ for all $x \in P$. So $b \in P^o$. But on the other hand $< b, y >> 1$ which contradicts the assumption $y \in (P^o)^o$.

2. We claim that if $P = conv(a_1, \ldots, a_m)$ then $P^o = \{x \in \mathbb{R}^d :< x, a_i >\leq 1\}$. Recall Lemma 1.0.2 says that any element of P is of the form $l_1a_1 + l_2a_2 + \ldots + l_ma_m$ with $l_i \geq 0$ and adding to 1. Hence if $< x, a_i >\leq 1$ for all i it is true for all of P.

Now we are ready to prove the Weyl-Minkowski theorem:
Proof: **Weyl-Minkowski** First we verify that a bounded polyhedron is a poly-

tope: Let P be $\{x \in \mathbb{R}^d :< x, c_i >\leq b_i\}$.

Consider the set of points E in P that are the unique intersection of d or more of the defining hyperplanes. The cardinality of E is at most $\binom{m}{d}$ so it is clearly a finite set and all its element are on the boundary of P. Denote by Q the convex hull of all elements of E. Clearly Q is a polytope and $Q \subset P$. We claim that $Q = P$. Suppose there is a $y \in P - Q$. Since Q is closed and bounded (bounded) we can find a linear functional f with the property that $f(y) > f(x)$ for all $x \in Q$. Now P is compact too, hence f attains its maximum on the boundary moreover we claim it must reach it in a point of E. The reason is that a boundary point that is not in E is in the solution set

We verify next that a polytope is indeed a polyhedron: We can assume that the polytope contains the origin in its interior (otherwise translate). So for a sufficiently small ball centered at the origin we have $B(0,r) \subset P$. Hence $P^o \subset B(0,r)^o = B(0,1/r)$. This implies that P^o is a bounded polyhedron. But we saw in the first part that bounded polyhedra are polytopes. Then P^o is a polytope. We are done because we know from the above lemma that $(P^o)^o = P$ and polar of polytopes are polyhedra.

1.0.2 Farkas, Carathéodory, and duality

When is a polytope non-empty? The key is that empty polyhedra come with some kind of "certificate" via the Farka's lemma. Now we present a simple proof of Farkas'lemma and then talk about some of its important consequences.

Theorem 1.0.6 *For a system of equations $Ax = b$, where A is a matrix and b is a vector. One and only one of the following choices holds:*

- *There is a non-negative vector x with $Ax = b$.*
- *There exists a non-trivial vector c such that $cA \geq 0$ but $c \cdot b < 0$.*

Proof: Clearly if the second option holds there cannot be positive solution for $Ax = b$ because it gives $0 \leq (cA)x = c(Ax) = cb < 0$.

Now suppose that $yb \geq 0$ for all y such that $yA \geq 0$. We want to prove that then b is an element of the cone K generated by the non-negative linear combinations of columns of A. For every b in $\mathbb{R}^n$ there exist in the cone $K = \{Ax | x \geq 0\}$ a point a that is closes to b and $Ax = a$ for $x \geq 0$. This observation is quite easy to prove and we leave it as an exercise (there are very easy arguments when the cone K is pointed). Now using this observation we have that

$$(A_j, b - a) \leq 0, \quad j = 1...k \quad (*)$$

and

$$(-a, b - a) \leq 0 \quad (**)$$

.

Why? the reason is a simple inequality. If we do not have the inequalities above we get for a sufficiently small $t \in (0, 1)$:

$$|b - (a + tA_j)|^2 = |(b - a) - tA_j|^2 =$$

$$|b-a|^2 - 2t(A_j, b-a) + t^2|A_j|^2 < |b-a|^2$$

or similarly we would get

$$|b-(a-ta)|^2 = |(b-a)+ta|^2 = |b-a|^2 - 2t(-a, b-a) + t^2|a|^2| < |b-a|^2$$

Both inequalities contradict the choice of a because $a + tA_j$ is in K and the same is true for $a - ta = (1-t)a \in K$. We have then that from the hypothesis and the equations in $(*)$ that $(b, -(b-a)) \geq 0$, which is the same as $(b, b-a) \leq 0$ and this together with equation (**) $(-a, b-a) \leq 0$ gives $(b-a, b-a) = 0$, and in consequence $b = a$. □

The theorem above is equivalent to this

Theorem 1.0.7 *For a system of inequalities $Ax \leq b$, where A is a matrix and b is a vector. One and only one of the following choices holds:*

- *There is a vector x with $Ax \leq b$.*
- *There exists a vector c such that $c \cdot b < 0$, $c \geq 0$, $\sum c_i > 0$, and $cA = 0$.*

The reason is simple, The system of inequalities $Ax \leq b$ has a solution if and only if for the matrix $A' = [I, A, -A]$ there is a non-negative solution to $A'x = b$. The rest is only a translation of the previous theorem in the second alternative.

Note that if you try to solve systems of **strict** inequalities, you would run into troubles for most computer programs. One needs to observe that for one strict inequality $bx < 0$ has a solution precisely when $bx \leq -1$ has a solution. Indeed, x has $bx = -q < 0$, if and only if $b(x/q) \leq -1$.

There is something really profound about Farkas's lemma. We can rephrase in the following way: The system of inequalities $Ax \leq b$ has a solution if and only if $yb \geq 0$, $yA = 0$, $y \geq 0$ has a solution. You have two systems that are match to each other. They have different number of variables and therefore represent polyhedra in different spaces and dimensions. But still they share this property. This phenomenon actually implies a similar matching phenomenon between a maximization LP and a minimization LP:

Theorem 1.0.8 (von Neumann's strong duality theorem of linear programming) *Let A be a matrix and b, c vectors (of adequate dimensions). Then*

$$max(cx: \ Ax \leq b) = min(yb : yA = c, \ y \geq 0)$$

,

provided both the primal and the dual polyhedra are non-empty.

The proof of this result follows from Farkas's lemma. Now we have an easy but important corollary:

Corollary 1.0.9 (Complementary Slackness) *Given dual linear problems*

$$max(cx: \ Ax \leq b) = min(yb : yA = c, \ y \geq 0)$$

, such that both have finite feasible solutions x_0, y_0 respectively, the following conditions are equivalent:

1. *x_0, y_0 are optimum solutions.*
2. *$cx_0 = y_0 b$*
3. *if a component of y_0 is positive, then the corresponding inequality in $Ax \leq b$ is satisfied by x_0 with equality.*

Proof: (1) and (2) are equivalent because of the strong duality theorem. (2) and (3) are equivalent because $cx_0 = y_0 Ax_0 = y_0 b$ happens if and only if $y_0(b - Ax_0) = 0$. Now since both y_0 and $b - Ax_0$ are non-negative, $y_0(b - Ax_0) = 0$ happens precisely when $b - Ax_0 = 0$ if $y_0 > 0$.

Lemma 1.0.10 (Carathéodory's theorem) *Let $X \subset \mathbb{R}^n$, and suppose $\bar{x} \in cone(X)$. Then $\bar{x} \in cone(x_1, \ldots, x_d)$ for some linearly independent vectors $x_1, \ldots, x_d \in X$. Thus $d \leq n$.*

Proof: Since $\bar{x} \in X$ there is at least one finite subset of vectors of which $\bar{x}$ is a non-negative linear combination. Take such a set $x_1, x_2, \ldots, x_m$ of smallest possible cardinality and such that $\bar{x} = \sum_{i=1}^{m} \alpha_i x_i$. We claim that it must be a linearly independent set of vectors. Suppose on the contrary that there is a linear dependence $\sum_{i=1}^{m} \gamma_i x_i = 0$ with some of the γ_i strictly positive and some strictly negative. Let $\tau = min(\alpha_i/\gamma_i : \ \gamma_i > 0)$. Then $\sum_{i=1}^{m}(\alpha_i - \tau\gamma_i)x_i = \bar{x}$. But by construction contains at least one less element. A contradiction.

Corollary 1.0.11 *If the optimum in the LP problems*

$$max(cx: \ Ax \leq b) = min(yb : yA = c, \ y \geq 0)$$

, is finite, then the minimum is attained at a vector $y_0 \geq 0$ whose positive components correspond to a set of linearly independent rows of the matrix A.

Proof: Suppose x^* is an optimal solution, thus $\delta = cx^* = min$. Therefore (c, δ) is a non-negative combination of the rows of A. By Carathéodory's lemma, there is exist $y_0 \geq 0$ whose positive components are linearly independent rows of the extended matrix $[A, b]$. Let $[A', b']$ be exactly those rows. We wish to proof that A' are linearly independent rows. But we know by complementary slackness that $A'x^* = b'$, thus A' has indeed full row rank.

1.0.3 The Faces of a Polytope.

Let P be a polytope in $\mathbb{R}^d$. A linear inequality $f(x) \leq \alpha$ is said to be *valid* on P if every point in P satisfies it. A set $F \subset P$ is a face of P if and only there exists a linear inequality $f(x) \leq \alpha$ which is valid on P and such that $F = \{x \in P : f(x) = \alpha\}$. In this case f is called a *supporting function* of F and the hyperplane defined by f is a *supporting hyperplane* of F. For a

face F consider the smallest affine subspace $aff(F)$ in $\mathbb{R}^d$ generated by F. Its dimension is called the dimension of F. Similarly we define the dimension of the polytope P, which we denote by $dim(P)$. An equivalent definition of dimension can be given in terms of affinely independent sets. We say the polyhedron P has dimension k if the largest collection of affinely independent vectors it contains has cardinality $k + 1$.

A face of dimension 0 is called a vertex. A face of dimension 1 is called an *edge*, and a face of dimension $dim(P) - 1$ is called a *facet*. The empty set is defined to be a face of P of dimension -1. Faces that are not the empty set or P itself are called proper.

Proposition 1.0.12 *Let $P = conv(a_1, \ldots, a_n)$ and $F \subset P$ a face. Then $F = conv(a_i, a_i \in F)$.*

Proof: Let $f(x) = \alpha$ be the supporting hyperplane. That $Q = conv(a_i, a_i \in F)$ is contained in F is clear. For the converse take $x \in F - Q$. We can still write x as $\lambda_1 a_1 + \ldots + \lambda_n a_n$ with the lambdas as usual. Applying f we get that if $\lambda_j > 0$ for an index not in Q, then we get $f(x) < \alpha$ because $f(a_j) < \alpha$ thus $\lambda_1 f(a_1) + \ldots + \lambda_n f(a_n) < \lambda_1 \alpha + \ldots \lambda_n \alpha = \alpha$. nd thus we arrive to a contradiction

Corollary 1.0.13 *A polytope has a finite number of faces, in particular a finite number of vertices and facets.*

For a polytope with vertex set $V = \{v_1, v_2, \ldots, v_n\}$ the *graph of P* is the abstract graph with vertex set V and the set of edges $E = \{(v_i, v_j) : [v_i, v_j] is\ an\ edge\ of\ P\}$. You can have a very entertaining day by drawing the graphs of polytopes. Later on we will prove a lot of cute properties about the graph of a polytope. Now there is a serious problem. We still don't have a formal verification that the graph of a polytope under our definition is non-empty! we must verify that there is always at least a vertex in a polytope. Such a seemingly obvious fact requires a proof. From looking at models of polyhedra one is certain that there is a containment relation among faces: a vertex of an edge that lies on the boundary of several facets, etc. Here is a first step to understand the

Lemma 1.0.14 *Let P be a d-polytope and $F \subset P$ be a face. Let $G \subset F$ be a face of F. Then G is a face of P as well.*

Proof: Suppose $P = conv(a_1, \ldots, a_n)$ and $f(x) = \alpha$ is a supporting linear functional to the face F. We saw $F = conv(a_i : i \in I_F)$, and by definition of being a face $f(a_i) = \alpha$ if $i \in I_F$ and $f(a_i) < \alpha$ otherwise. At the same time we have $g(x) = \beta$ such that for $G = conv(a_j : j \in I_G \subset I_F)$. We have again that $g(a_i) = \beta$ if $i \in I_G \subset I_F$ and $g(x) < \beta$ otherwise.

Construct the hyperplane $h = f + \epsilon g$. Note that for a_j with $j \in I_G$ we have $h(a_i) = \alpha + \epsilon\beta$. Now for a_j with $j \in I_F - I_G$ we have $h(a_i) < \alpha + \epsilon\beta$. For all the other indices we have, by choosing ϵ small enough, we have that $h(a_i) < \alpha + \epsilon\beta$ too.

Corollary 1.0.15 *Every non-empty polytope has at least one vertex.*

Proof: Every non-empty polytope has a proper face. In fact one can be found as the maximum or minimum of a linear functional. If the polytope is not a point, this proper face has a smaller dimension. The face itself is a polytope which in turn has a proper face. Going this way we come to a point, which is a vertex.

Theorem 1.0.16 *Every polytope is the convex hull of the set of its vertices.*

Let P be a polytope and Q be the convex hull of its vertices. We now from the previous corollary that Q is not empty. Suppose P and Q are different. A point s in $P - Q$ we can separate s from Q by a linear function. Now that linear function attains a maximum in P that is again a vertex!! contradiction because it was not in Q.

Now we know that a polytope has a canonical representation as the convex hull of its vertices. The results above establishes that the set of all faces of a polytope form a partially ordered set by the order given by containment. This poset receives the name of the *face poset* of a polytope. We say that two polytopes are *combinatorially equivalent* or *combinatorially isomorphic* if their face posets are the same. In particular, two polytopes P, Q are isomorphic if they have the same number of vertices and there is a one-to-one correspondence p_i to q_i between the vertices such that $conv(p_i : i \in I)$ is a face of P if and only if $conv(q_i : i \in I)$ is a face of Q. The bijection is called an isomorphism.

Another property that can be extrapolated from looking at the Platonic solids is that there is a duality relation where two polytopes are matched to each other by paring the vertices of one with the facets of the other and vice versa. We want now to make this intuition precise. We will establish a bijection

between the faces of P and the faces of P^o. Let $P \subset \mathbb{R}^d$ be a d-dimensional polytope containing the origin as its interior point. For a non-empty face F of P define

$$\hat{F} = \{x \in P^o :< x, y >= 1 for\ all\ y \in F\}$$

and for the empty face define $\hat{} = Q$.

Theorem 1.0.17 *The hat operation applied to faces of a d-polytope P satisfies*

1. *The set $\hat{F}$ is a face of P^o*

2. $dim(F) + dim(\hat{F}) = d - 1$.

3. *The hat operation is involutory: $(\hat{\hat{F}}) = F$.*

4. *If $F, G \subset P$ are faces and $F \subset G \subset P$, then $\hat{G}, \hat{F}$ are faces of P^o and $\hat{G} \subset \hat{F}$.*

Proof: To set up notation we take $P = conv(a_1, a_2, \ldots, a_m)$ and $F = conv(a_i : i \in I)$.
(1) Define $v := 1/|I| \sum_{i \in I} a_i$. We claim that in fact, $\hat{F} = \{x \in P^o :< x, v >= 1\}$. It is clear that $\hat{F} \subset \{x \in P^o :< x, v >= 1\}$ The reasons for the other containment are: we already know that $< x, a_i > \leq 1$ and $< x, v >= 1$ implies then that $< x, a_i >= 1$ for all $i \in I$. Since all other elements of F are convex linear combinations of a_i's we are done.

Now that the set $\hat{F}$ is a face of P^o is clear because the supporting hyperplane to the face is the linear functional $< x, v >= 1$. Warning! the $\hat{F}$ could be still empty face!!
(2) Now we convince ourselves that if F is a non-empty face, then $\hat{F}$ is non-empty and moreover the sum of their dimensions is equal to $d - 1$.

By definition of face $F = \{x :< x, c >= \alpha\}$ and for other points in P we have $< y, c >< \alpha$. Because the origin is in P we have that $\alpha > 0$, which means that $b = c/\alpha \in \hat{F}$ because 1) $< b, a_i >= 1$ for $i \in I$ and 2) $< b, a_i > \leq 1$ (this second observation is a reality check: b is in P^o). Hence $\hat{F}$ is not empty.

Suppose $dim(F) = k$ and let $h_1, \ldots, h_{d-k-1} \in \mathbb{R}^d$ be linear independent vectors orthogonal to the linear span of F. The orthogonality means that $< h_i, a_j >= 0$ for $j \in I$ and all h_i. We complete to a basis!

For all sufficiently small values $\epsilon_1, \ldots, \epsilon_{d-k-1}$ we have that $r := b + \epsilon_1 h_1 + \epsilon_2 h_2 + \ldots + \epsilon_{d-k-1} h_{d-k-1}$ satisfies $< r, a_i >= 1$ for $i \in I$ and $< r, a_j >< 1$ for other indices. Hence r is in $\hat{F}$ proving that $dim(\hat{F}) \geq d - 1 - dim(F)$.

On the other hand $\hat{F}$ is in the intersection of the hyperplanes $\{x \in \mathbb{R}^d :< x, a_i >= 1\}$ therefore $dim(\hat{F}) \leq d - 1 - dim(F)$. We are done.
(3) Denote by G the set $\{x \in P :< x, y >= 1 for\ y \in P\}$. We know from the previous two parts that G is a face of P and it has the same dimension as F!! and F is contained in G! who else can it be? why?
(4) Suppose again that $F \subset G$. If $x \in \hat{G}$ then $< x, y >= 1$ for all $y \in G$ in particular for all members of F hence $x \in \hat{F}$.

From now on, by the *dual* of a polytope $P \subset \mathbb{R}^d$ we mean the following: Consider the smallest affine subspace containing P, so that P has full dimension.

Move the origin to be inside P and apply the polar operation to P. Regardless of where exactly you put the origin you will get the same combinatorial type of polytope!! thus if we are only interested on the combinatorial structure of polytopes the dual will be unique.

Graphs of polytopes have been studied because of the connection to the theory of linear programming. Several open questions about their diameter remain. In the 1960's Balinski proved a now classical theorem about the graphs of polytopes:

Theorem 1.0.18 *The graphs of d-dimensional polytopes are d-connected.*

before we start the proof we should observe that given a linear function $l(x) : \mathbb{R}^d \to \mathbb{R}$ such that $l(v_i) \neq l(v_j)$ for any pair of vertices of a d-polytope P, we can orient or direct the edges of P from v_i to v_j whenever $l(v_i) < l(v_j)$ and v_i, v_j are adjacent. The orientation we have produced allows you to find the "winning" vertex that maximizes the linear function by simply walking along the graph!! the basic reason is the following lemma:

Lemma 1.0.19 *For a d-polytope P and a linear function l that is not constant in an edge of P. Then for every vertex v_0 of P either $l(v) = max\{l(x) : x \in P\}$ or there exists a neighbor u of v such that $l(u) > l(v)$.*

Proof: Denote by $v_1, v_2, \ldots, v_k$ the neighbors of v_0 in the graph of P. It is enough to be sure that any point of P, say x, can be written in the form

$$x = v_0 + \sum_{i=1}^{k} \lambda(v_i - v_0).$$

With positive coefficients. The reason this is enough is because if one such expression exists $l(v_i - v_0)$ is positive precisely when moving from v_0 to v_i increases the value of l. If no $l(v_i - v_0)$ is positive then clearly staying at v_0 is best you can do. How can we prove that equation above? Take the interval segment $[x, v_0]$. For a point r inside the interval close enough to v_0 we have that indeed

$r = t(x - v_0) = \sum_{i=1}^{k} s_i(v_i - v_0)$

By the property of being in a convex hull. So we are done. □

Now we prove Balinski's theorem:

Proof: **(Balinski's theorem)** We proceed by induction on the dimension of the polytope. It is clear the theorem holds for dimension one because polytopes are simply segments. Assume the theorem is true for all polytopes of dimension less than d and suppose you have a d-polytope with a vertex-cut $v_1, \ldots, v_s$ with $s < d$. This means that $s + 1 \leq d$. Because in $\mathbb{R}^d$ a hyperplane is determined by d points we can make a H hyperplane pass through the v_i's plus at least one vertex not in the set.

There are two cases: 1) H is supporting hyperplane for a face or 2) H cuts through the interior of the polytope. In the first case find the hyperplane H' that parallel to H that supports an another face. For any two points not in the cut we can move, using the previous lemma, until we reach H'. There by induction any two points are connected, so we have find a connecting path.

In the second case. We again have two hyperplanes H' and H'' parallel to H supporting the polytope from below and from above in two faces F', F''. if the two points not in the cut are in the same side of H we repeat the argument to connect via the use of F' or F'' depending on which side the points are present. If they are in opposite sides we still have the extra point along H to connect them (we can construct an increasing path from that extra point to F' or a decreasing path to F''). □