Math 583E: Linear and Integer Polyhedra

Rekha R. Thomas

Department of Mathematics, University of Washington, Seattle, WA 98195
E-mail address: thomas@math.washington.edu

Abstract. The material in these notes is drawn from several existing sources, among which the main ones are the book *Theory of Linear and Integer Programming* by Alexander Schrijver [**Sch86**] and an unpublished set of lecture notes by Les Trotter titled *Lectures on Integer Programming* [**Tro92**].

Please send all corrections and suggestions, no matter how trivial, to thomas@math.washington.edu

Contents

CHAPTER 1

Farkas Lemma

The Farkas Lemma is sometimes called the **Fundamental Theorem of Linear Inequalities** and is due to Farkas and Minkowski with sharpenings by Caratheodory and Weyl. It underlies all of *linear programming* and is the first of several *alternative theorems* that we will see in this course. It is can be interpreted as a theorem about *polyhedral cones* which makes it a geometric statement and therefore, easier to remember.

DEFINITION 1.1. (1) A non-empty set $C \subseteq \mathbb{R}^n$ is a **cone** if for every $\mathbf{x}, \mathbf{y} \in C$, $\lambda\mathbf{x} + \mu\mathbf{y}$ is also in C whenever $\lambda, \mu \in \mathbb{R}_{\geq 0}$.

(2) A cone C is **polyhedral** if there exists a matrix $A \in \mathbb{R}^{m \times n}$ such that

$$C = \{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} \leq \mathbf{0}\}.$$

Such a cone is said to be *finitely constrained*.

(3) The cone generated by the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathbb{R}^n$ is the set

$$\mathrm{cone}(\mathbf{b}_1, \ldots, \mathbf{b}_p) := \{\sum_{i=1}^{p} \lambda_i \mathbf{b}_i \,:\, \lambda_i \geq 0\} = \{\lambda^t B \,:\, \lambda \geq \mathbf{0}\}$$

where B is the matrix with rows $\mathbf{b}_1, \ldots, \mathbf{b}_p$ and $\lambda = (\lambda_i)$. Such a cone is said to be *finitely generated*.

One can check that a cone C is a **convex set** (i.e., if $\mathbf{x}, \mathbf{y} \in C$ then $\lambda\mathbf{x} + \mu\mathbf{y} \in C$ for all $0 \leq \lambda, \mu \leq 1$ and $\lambda + \mu = 1$.) For a non-zero vector $\mathbf{a} \in \mathbb{R}^n$ we call the set $\{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{a}\mathbf{x} \leq 0\}$ a **linear half-space** and the set $\{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{a}\mathbf{x} \leq \beta\}$ for some scalar $\beta \neq 0$, an **affine half-space**. Thus a polyhedral cone is the intersection of finitely many linear half-spaces.

Our first goal is to outline a classical elimination scheme for linear inequalities called **Fourier-Motzkin Elimination**. Suppose $A \in \mathbb{R}^{m \times n}$ with rows $\mathbf{a}_1, \ldots, \mathbf{a}_m$ and $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{R}^m$. Consider the inequality system

$$A\mathbf{x} \leq \mathbf{b}. \tag{1}$$

The Fourier-Motzkin procedure eliminates x_n from (1) to get a new system (2) that will have the property that (1) is feasible if and only if (2) is feasible.

To create (2) we do the following operations to the inequalities in (1):

(i) If $a_{in} = 0$ then put $\mathbf{a}_i\mathbf{x} \leq b_i$ in (2).

(ii) For every pair of rows $\mathbf{a}_i$ and $\mathbf{a}_j$ of A such that $a_{in} > 0$ and $a_{jn} < 0$ put the following inequality in (2).

$$\begin{array}{rlllll}
 & \frac{a_{i1}}{a_{in}}x_1 & +\cdots+ & \frac{a_{in-1}}{a_{in-1}}x_{n-1} & + \frac{a_{in}}{a_{in}}x_n & \leq \frac{b_i}{a_{in}} \\[2ex]
- & \frac{a_{j1}}{a_{jn}}x_1 & -\cdots- & \frac{a_{jn-1}}{a_{jn-1}}x_{n-1} & - \frac{a_{jn}}{a_{jn}}x_n & \leq -\frac{b_j}{a_{jn}} \\ \hline
 & (\frac{a_{i1}}{a_{in}} - \frac{a_{j1}}{a_{jn}})x_1 & +\cdots+ & (\frac{a_{in-1}}{a_{in}} - \frac{a_{jn-1}}{a_{jn}})x_{n-1} & & \leq \frac{b_i}{a_{in}} - \frac{b_j}{a_{jn}}
\end{array}$$

Lemma 1.2. *The system (1) is feasible if and only if the system (2) constructed as above is feasible.*

Proof. If $\mathbf{x} = (x_1, \ldots, x_n)$ satisfies (1), then $(x_1, \ldots, x_{n-1})$ satisfies (2) since all inequalities in (2) are non-negative linear combinations of those in (1).

Now suppose $(x_1, \ldots, x_{n-1})$ satisfies (2). We need to construct an x_n such that $\mathbf{x} = (x_1, \ldots, x_{n-1}, x_n)$ satisfies (1). Define x_n as follows: Rearranging the inequality obtained from (ii) we have that for all $a_{in} > 0$ and $a_{jn} < 0$, $(x_1, \ldots, x_{n-1})$ satisfies

$$\frac{1}{a_{in}}(b_i - a_{i1}x_1 - \cdots - a_{in-1}x_{n-1}) \geq \frac{1}{a_{jn}}(b_j - a_{j1}x_1 - \cdots - a_{jn-1}x_{n-1}).$$

Therefore, there exists λ and μ such that:

$$\lambda := \min_{(i\,:\,a_{in}>0)}\{\frac{1}{a_{in}}(b_i - a_{i1}x_1 - \cdots - a_{in-1}x_{n-1})\}, \text{ and}$$

$$\mu := \max_{(j\,:\,a_{jn}<0)}\{\frac{1}{a_{jn}}(b_j - a_{j1}x_1 - \cdots - a_{jn-1}x_{n-1})\}.$$

If there does not exist any $\mathbf{a}_i$ such that $a_{in} > 0$ then set $\lambda = \infty$ and similarly, if there does not exist any $\mathbf{a}_j$ such that $a_{jn} < 0$, set $\mu = -\infty$. Now if x_n is chosen such that $\lambda \geq x_n \geq \mu$, then $\mathbf{x}$ satisfies (1) since

$$\frac{1}{a_{in}}(b_i - a_{i1}x_1 - \cdots - a_{in-1}x_{n-1}) \geq x_n \geq \frac{1}{a_{jn}}(b_j - a_{j1}x_1 - \cdots - a_{jn-1}x_{n-1})$$

whenever $a_{in} > 0$ and $a_{jn} < 0$. Trivially, the inequalities from (i) are satisfied by $\mathbf{x}$. □

Remark 1.3. (1) Note that if the matrix A and vector $\mathbf{b}$ are rational in the system $A\mathbf{x} \leq \mathbf{b}$, then the system (2) obtained by eliminating x_n is also rational.

(2) Geometrically, eliminating x_n is equivalent to projecting the solution set of $A\mathbf{x} \leq \mathbf{b}$ onto the $x_1, \ldots, x_{n-1}$ coordinates. This follows from the proof of Lemma 1.2 since the proof showed that $(x_1, \ldots, x_n)$ satisfies (1) if and only if $(x_1, \ldots, x_{n-1})$ satisfies (2).

(3) Suppose in each row $\mathbf{a}_i$ of A, $a_{in} \geq 0$. Then step (ii) in the Fourier-Motzkin elimination procedure does not contribute any inequalities to (2). If $(x_1, \ldots, x_{n-1})$ satisfies (2) (which consists of the inequalities from (1) in which x_n was absent) then choose

$$x_n = \min_{\{j\,:\,a_{jn}>0\}} \frac{b_j - \sum_{k=1}^{n-1} a_{jk}x_k}{a_{jn}}.$$

Then $(x_1, \ldots, x_n)$ satisfies (1). A similar argument works if $a_{in} \leq 0$ for all i. In particular, if for all rows $\mathbf{a}_i$ of A, $a_{in} > 0$, then (2) is a vacuous system of inequalities in $x_1, \dots, x_{n-1}$ which means that every point in $\mathbb{R}^{n-1}$ lifts to a point that satisfies (1).

(4) Fourier-Motzkin elimination allows the solution of the system $A\mathbf{x} \leq \mathbf{b}$.

(5) The Fourier-Motzkin method does not guarantee that the number of inequalities created in its intermediate steps is polynomial in the number of inequalities in the initial system. An example of this is given on pp. 156 of [**Sch86**].

Let p be a positive integer and $n := 2^p + p + 2$. Take any linear system $A\mathbf{x} \leq \mathbf{b}$ in n variables that contains in its left-hand-side the $8\binom{n}{3}$ possible linear forms

$$\pm x_i \pm x_j \pm x_k, \;\; 1 \leq i < j < k \leq n.$$

It can be shown (by induction) that after eliminating the first t variables the following linear forms are in the left-hand-side of the resulting system:

$$\pm x_{j_1} \pm x_{j_2} \pm \cdots \pm x_{j_s}, \;\; t+1 \leq j_1 < j_2 < \cdots < j_s \leq n$$

where $s := 2^t + 2$. Therefore, after p eliminations (recall that $n := 2^p + p + 2$) there are at least $2^{2^p+2} = 2^{n-p} \geq 2^{\frac{n}{2}}$ inequalities.

Exercise 1.4. Find an example of an inequality system with m inequalities in n variables that results, after projecting out some set of the variables, in a system that has t *non-redundant* inequalities where t is not polynomial in m and n. (Can you make the example in Remark 1.3 (5) have this property?)

Theorem 1.5. **Weyl's Theorem.** *If a non-empty cone C is finitely generated, then it is also finitely constrained, or equivalently, polyhedral.*

Proof. Let $C = \{\lambda B \,:\, \lambda \geq \mathbf{0}\} \subseteq \mathbb{R}^n$ be a finitely generated cone with $B \in \mathbb{R}^{p \times n}$. Then

$$\begin{aligned} C &= \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{x} = \lambda B, \lambda \geq \mathbf{0}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{x} - \lambda B \leq \mathbf{0}, -\mathbf{x} + \lambda B \leq \mathbf{0}, -\lambda \leq \mathbf{0}\} \\ &= \{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} \leq \mathbf{0}\} \end{aligned}$$

where $A\mathbf{x} \leq \mathbf{0}$ is obtained from the inequality system $\mathbf{x} - \lambda B \leq \mathbf{0}, -\mathbf{x} + \lambda B \leq \mathbf{0}, -\lambda \leq \mathbf{0}$ by Fourier-Motzkin elimination of λ. $\square$

Theorem 1.6. **Farkas Lemma.** *Given a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{b} \in \mathbb{R}^n$, exactly one of the following holds:*

either $\mathbf{y}A = \mathbf{b}$, $\mathbf{y} \geq \mathbf{0}$ has a solution or there exists an $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} \leq \mathbf{0}$ but $\mathbf{bx} > 0$.

Proof. We first check that both statements cannot hold simultaneously. If there exists a $\mathbf{y}^* \geq \mathbf{0}$ such that $\mathbf{y}^* A = \mathbf{b}$ and an $\mathbf{x}^*$ such that $A\mathbf{x}^* \leq \mathbf{0}$ and $\mathbf{bx}^* > 0$ then

$$0 < \mathbf{bx}^* = (\mathbf{y}^* A)\mathbf{x}^* = \mathbf{y}^*(A\mathbf{x}^*) \leq 0$$

which is a contradiction. Let $C = \{\mathbf{y}A \,:\, \mathbf{y} \geq \mathbf{0}\}$. By Theorem 1.5, there exists a matrix B such that $C = \{\mathbf{x} \,:\, B\mathbf{x} \leq \mathbf{0}\}$. Now $\mathbf{b} = \mathbf{y}A,\; \mathbf{y} \geq \mathbf{0}$ if and only if $\mathbf{b} \in C$. On the other hand, $\mathbf{b} \notin C$ if and only if there exists a row $\mathbf{x}$ of B such that $\mathbf{bx} > 0$. Since every row of A lies in C (by using $\mathbf{y}$ equal to the unit vectors) , $A\mathbf{x} \leq \mathbf{0}$. Therefore $\mathbf{b} \notin C$ if and only if $A\mathbf{x} \leq \mathbf{0},\; \mathbf{bx} > 0$ is feasible. $\square$

Definition 1.7. The **polar** of a cone $C \subseteq \mathbb{R}^n$ is the set

$$C^* := \{\mathbf{z} \in \mathbb{R}^n \,:\, \mathbf{z} \cdot \mathbf{x} \leq 0 \;\; \forall \;\; \mathbf{x} \in C\}.$$

EXERCISE 1.8. (**HW**) Prove the following facts for cones C and K:

(1) $C \subseteq K$ implies that $C^* \supseteq K^*$,
(2) $C \subseteq C^{**}$,
(3) $C^* = C^{***}$,
(4) If $C = \{\lambda B : \lambda \geq \mathbf{0}\}$ then $C^* = \{\mathbf{x} : B\mathbf{x} \leq \mathbf{0}\}$,
(5) If $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$ then $C = C^{**}$.

THEOREM 1.9. **Minkowski's Theorem.** *If a cone C is polyhedral then it is non-empty and finitely generated.*

PROOF. Let $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$. Then, since $\mathbf{x} = \mathbf{0}$ lies in C, $C \neq \emptyset$. Let $L := \{\lambda A : \lambda \geq \mathbf{0}\}$. By Exercise 1.8 (4), $C = L^*$ and hence, $C^* = L^{**}$. Since L is finitely generated, by Theorem 1.5, L is also polyhedral which implies by Exercise 1.8 (5) that $L = L^{**}$ and hence $C^* = L$. Thus C^* is finitely generated. Now since L is finitely generated, by Theorem 1.5, L is also polyhedral. So repeating the same arguments as above for L, we get that L^* is finitely generated. But we saw that $L^* = C$ and hence C is finitely generated. $\square$

Combining Theorems 1.5 and 1.9 we get the the **Weyl-Minkowski duality** for cones:

THEOREM 1.10. *A cone is polyhedral if and only if it is finitely generated.*

Farkas Lemma provides analogs of Exercise 1.8 (4) and (5).

COROLLARY 1.11. (1) *If $C = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}$ then $C = C^{**}$.*
(2) *If $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$ then $C^* = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}$.*

PROOF. (1) Farkas Lemma states that either there exists $\mathbf{y} \geq \mathbf{0}$ such that $\mathbf{b} = \mathbf{y}A$ or there exists $\mathbf{x}$ such that $A\mathbf{x} \leq \mathbf{0}$ and $\mathbf{b}\mathbf{x} > 0$. Let $C = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}$. By Exercise 1.8 (4), $C^* = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$. The first statement of Farkas Lemma says that $\mathbf{b} \in C$. The second statement states that there exists $\mathbf{x} \in C^*$ such that $\mathbf{b}\mathbf{x} > 0$, or equivalently, $\mathbf{b} \notin C^{**}$. Further, the two statements are mutually exclusive. Therefore, $\mathbf{b} \notin C$ if and only if $\mathbf{b} \notin C^{**}$ which implies that $C = C^{**}$.

(2) If $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$, then by Exercise 1.8 (4), $C = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}^*$. Therefore by part (1), $C^* = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}^{**} = \{\mathbf{y}A : \mathbf{y} \geq \mathbf{0}\}$. $\square$

EXERCISE 1.12. The **Fundamental Theorem of Linear Equations** says that given a matrix A and a vector $\mathbf{b}$ exactly one of the following holds:

either $A\mathbf{x} = \mathbf{b}$ has a solution or there exists a vector $\mathbf{y}$ such that $\mathbf{y}A = \mathbf{0}$ but $\mathbf{y}\mathbf{b} \neq 0$.

Prove this theorem.

EXERCISE 1.13. (**HW**) Prove the following two variants of Farkas Lemma.

(1) Given a A and a vector $\mathbf{b}$ exactly one of the following holds. Either the system $A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$ has a solution or there exists $\mathbf{y}$ such that $\mathbf{y}A \leq \mathbf{0}$ but $\mathbf{y}\mathbf{b} > 0$.
(2) Given a matrix A and a vector $\mathbf{b}$ exactly one of the following holds. Either the system $A\mathbf{x} \leq \mathbf{b}$ has a solution or there exists $\mathbf{y} \leq \mathbf{0}$ such that $\mathbf{y}A = \mathbf{0}$ but $\mathbf{y}\mathbf{b} > 0$.

EXERCISE 1.14. (**HW**) This exercise will provide an alternate algorithm for computing projections of feasible regions of inequality systems.

(1) Suppose $Q = \{(\mathbf{u}, \mathbf{x}) \in \mathbb{R}^p \times \mathbb{R}^q : A\mathbf{u} + B\mathbf{x} \leq \mathbf{b}\}$ where $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{m \times q}$. Let

$$\mathrm{Proj}_{\mathbf{x}}(Q) := \{\mathbf{x} \in \mathbb{R}^q : \exists \mathbf{u} \in \mathbb{R}^p : (\mathbf{u}, \mathbf{x}) \in Q\}$$

be the projection of Q into $\mathbb{R}^q$. By Fourier-Motzkin elimination we know that there exists a $C \in \mathbb{R}^{m' \times q}$ and a vector $\mathbf{d} \in \mathbb{R}^{m'}$ such that $\mathrm{Proj}_{\mathbf{x}}(Q) = \{\mathbf{x} \in \mathbb{R}^q : C\mathbf{x} \leq \mathbf{d}\}$. We could compute C and $\mathbf{d}$ using Fourier-Motzkin elimination. But here is an algorithm that uses the Weyl-Minkowski theorem for cones.

Let $W := \{\mathbf{v} : \mathbf{v}A = 0,\ \mathbf{v} \geq 0\}$. Note that W is a cone. Let $\mathcal{V}(W)$ be a set of generators for the extreme rays of W. Prove that

$$\mathrm{Proj}_{\mathbf{x}}(Q) = \{\mathbf{x} \in \mathbb{R}^q : (\mathbf{v}B)\mathbf{x} \leq \mathbf{v}\mathbf{b},\ \mathbf{v} \in \mathcal{V}(W)\}.$$

(2) The above exercise allows projections of inequality systems to be computed via algorithms that will transform one representation of a cone to the other. You could also do this projection using a software package such as PORTA [**CL**] that does Fourier-Motzkin elimination. Do the following projection using the algorithm outlined in (1) and then double check against what **fmel** gives you in Porta.

Eliminate the variables y_2, y_3, y_4 from the following linear inequality system.

$$\begin{array}{cccccccccc}
y_2 & & & & & & & \leq & 0 \\
 & y_3 & & & & & & \leq & 0 \\
 & & y_4 & & & & & \leq & 0 \\
y_2 & +y_3 & & -x_1 & & & & \leq & 0 \\
y_2 & & +y_4 & -x_1 & & & & \leq & 0 \\
 & y_3 & +y_4 & -x_1 & & & & \leq & 0 \\
 & & & -x_1 & & & & \leq & 0 \\
-y_2 & & & & & & & \leq & 0 \\
 & -y_3 & & & & & & \leq & 0 \\
 & & -y_4 & & & & & \leq & 0 \\
-y_2 & & & +x_1 & +x_2 & & & \leq & 1 \\
 & -y_3 & & +x_1 & & +x_3 & & \leq & 1 \\
 & & -y_4 & +x_1 & & & +x_4 & \leq & 1 \\
-y_2 & -y_3 & & +x_1 & +x_2 & +x_3 & & \leq & 1 \\
-y_2 & & -y_4 & +x_1 & +x_2 & & +x_4 & \leq & 1 \\
 & -y_3 & -y_4 & +x_1 & & +x_3 & +x_4 & \leq & 1 \\
y_2 & & & & -x_2 & & & \leq & 0 \\
 & y_3 & & & & -x_3 & & \leq & 0 \\
 & & y_4 & & & & -x_4 & \leq & 0
\end{array}$$

The relevance of this system will be clear later. For now it's just an exercise.

CHAPTER 2

Polyhedra

Definition 2.1. If $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbb{R}^n$, we call $\sum_{i=1}^p \lambda_i \mathbf{a}_i$, $\sum_{i=1}^p \lambda_i = 1$, $\lambda_i \geq 0$ for all i, a **convex combination** of $\mathbf{a}_1, \ldots, \mathbf{a}_p$. The set of all convex combinations of finitely many points from a set $S \subseteq \mathbb{R}^n$, denoted as $\mathrm{conv}(S)$, is called the **convex hull** of S.

Definition 2.2.
(1) A set $P \subseteq \mathbb{R}^n$ of the form $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ for a matrix $A \in \mathbb{R}^{m \times n}$ and vector $\mathbf{b} \in \mathbb{R}^m$ is called a **polyhedron**.
(2) The convex hull of finitely many points in $\mathbb{R}^n$ is called a **polytope**.

Theorem 2.3. **Affine Weyl Theorem.** *If $P = \{\mathbf{y}B + \mathbf{z}C : \mathbf{y}, \mathbf{z} \geq \mathbf{0}, \sum z_i = 1\}$ for $B \in \mathbb{R}^{p \times n}$ and $C \in \mathbb{R}^{q \times n}$ then there exists a matrix $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$ such that $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$.*

Proof. If $P = \emptyset$, then we may choose $A = (0\ 0\ \cdots\ 0)$ and $b = -1$. (The case $P = \emptyset$ corresponds to B and C being vacuous.) If $P \neq \emptyset$ but C is vacuous, then the above theorem is Weyl's Theorem for cones (Theorem 1.5). So assume that $P \neq \emptyset$ and C is not vacuous. Define the finitely generated cone $P' \subseteq \mathbb{R}^{n+1}$ as follows.

$$P' := \{(\mathbf{y}, \mathbf{z}) \begin{pmatrix} B & \mathbf{0} \\ C & \mathbf{1} \end{pmatrix} : \mathbf{y}, \mathbf{z} \geq \mathbf{0}\}$$

where $\mathbf{1}$ is the vector of all ones of length q and $\mathbf{0}$ has length p. Then $P' = \{(\mathbf{y}B + \mathbf{z}C, \sum z_i) : \mathbf{y}, \mathbf{z} \geq \mathbf{0}\}$ is a finitely generated cone. Note that $\mathbf{x} \in P$ if and only if $(\mathbf{x}, 1) \in P'$. Applying Weyl's Theorem to P' we get

$$P' = \{(\mathbf{x}, x_{n+1}) : (A| - \mathbf{b}) \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \leq \mathbf{0}\}$$

for some matrix $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and some m. Since $\mathbf{x} \in P$ if and only if $(\mathbf{x}, 1) \in P'$, we get that $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$. $\square$

Definition 2.4. If P and Q are two sets in $\mathbb{R}^n$ then

$$P + Q = \{\mathbf{p} + \mathbf{q} : \mathbf{p} \in P,\ \mathbf{q} \in Q\} \subseteq \mathbb{R}^n$$

is called the **Minkowski sum** of P and Q.

Remark 2.5. Note that Theorem 2.3 says that the Minkowski sum of a polyhedral (finitely generated) cone and a polytope is a polyhedron.

We now prove that, in fact, every polyhedron is the Minkowski sum of a finitely generated cone and a polytope.

Theorem 2.6. **Affine Minkowski Theorem.** *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ for some $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then there exists $B \in \mathbb{R}^{p \times n}$ and $C \in \mathbb{R}^{q \times n}$ such that $P = \{\mathbf{y}B + \mathbf{z}C : \mathbf{y}, \mathbf{z} \geq \mathbf{0}, \sum z_i = 1\}$.*

Proof. If $P = \emptyset$ then we can take B and C to be vacuous (i.e., $p = q = 0$). If $P \neq \emptyset$ then define the polyhedral cone $P' \subseteq \mathbb{R}^{n+1}$ as follows.

$$P' := \{(\mathbf{x}, x_{n+1}) : \begin{pmatrix} A & -\mathbf{b} \\ \mathbf{0} & -1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \leq \mathbf{0}\}.$$

Again, $\mathbf{x} \in P$ if and only if $(\mathbf{x}, 1) \in P'$. By Minkowski's Theorem for cones, there exists a $D \in \mathbb{R}^{r \times (n+1)}$ such that

$$P' = \{\mathbf{u}D : \mathbf{u} \geq \mathbf{0}\}.$$

The rows of D are the generators of the cone P'. We may rearrange the rows of D so that the rows with last component zero are on the top. For each of the remaining rows of D, note that the last component is positive by noting the form of the last inequality in the constraint representation of P'. Therefore, we may rescale all these rows of D so that the last component in each row in one. This only rescales the generating set of P' by positive scalars. Hence we may assume that D is of the form

$$D = \begin{pmatrix} B & \mathbf{0} \\ C & \mathbf{1} \end{pmatrix}$$

where $B \in \mathbb{R}^{p \times n}$, $C \in \mathbb{R}^{q \times n}$ and $p + q = r$. Now $(\mathbf{x}, 1) \in P'$ if and only if $\mathbf{x} = \mathbf{y}B + \mathbf{z}C$ with $\mathbf{y}, \mathbf{z} \geq \mathbf{0}$, $\mathbf{u} = (\mathbf{y}, \mathbf{z})$ and $\sum z_i = 1$. $\square$

We proved both the affine Weyl and Minkowski theorems for polyhedra by a technique called *homogenization* which lifts polyhedra to cones in one higher dimension and then uses a theorem for cones followed by a projection back to the original space of the polyhedron. The adjective "homogenous" refers to the right-hand-side $\mathbf{0}$ in the constraint representation of a polyhedral cone. This is a standard technique for proving facts about polyhedra.

Exercise 2.7. Prove that a set $P \subseteq \mathbb{R}^n$ is a polytope if and only if P is a bounded polyhedron.

Example 2.8. The following is an example of converting from one representation of a polyedron to another using Porta. In this example, we wish to compute the convex hull of all points in $\mathbb{N}^4$ obtained by permuting the components of the vector $(1, 2, 3, 4)$. The resulting convex hull is called a **permutahedron** in $\mathbb{R}^4$. The input file should be called **filename.poi** and looks as follows.

```
DIM = 4

CONV_SECTION
1 2 3 4
1 2 4 3
1 3 2 4
1 3 4 2
1 4 2 3
1 4 3 2
2 1 3 4
2 1 4 3
2 3 1 4
2 3 4 1
2 4 1 3
```

```
2 4 3 1
3 1 2 4
3 1 4 2
3 2 1 4
3 2 4 1
3 4 1 2
3 4 2 1
4 1 2 3
4 1 3 2
4 2 1 3
4 2 3 1
4 3 1 2
4 3 2 1
END
```

In Porta we use the command **traf filename.poi** to get the following output which is written to the file **filename.poi.ieq**.

```
DIM = 4

VALID
4 3 2 1

INEQUALITIES_SECTION
(  1) +x1+x2+x3+x4 == 10

(  1) -x2-x3-x4 <= -6
(  2) -x2-x3    <= -3
(  3) -x2   -x4 <= -3
(  4)    -x3-x4 <= -3
(  5) -x2       <= -1
(  6)    -x3    <= -1
(  7)       -x4 <= -1
(  8)       +x4 <=  4
(  9)    +x3    <=  4
( 10) +x2       <=  4
( 11)    +x3+x4 <=  7
( 12) +x2   +x4 <=  7
( 13) +x2+x3    <=  7
( 14) +x2+x3+x4 <=  9

END
```

```
strong validity table :
\ P       |                              |
 \ O      |                              |
I \ I     |                              |
 N \ N    | 1     6     11    16    21   | #
  E \ T   |                              |
   Q \ S  |                              |
    S \   |                              |
       \  |                              |
--------------------------------------------
1         | ..... ..... ..... ...** **** :  6
2         | ..... ..... ..*.* ...*. *... :  4
3         | ..... ..... ...*. *...* .*.. :  4
4         | ..... ..... ..... .**.. ..** :  4
5         | ..... .**.. ..**. ...** .... :  6
6         | ..... ...*. *...* .*... *.*. :  6
7         | ..... ....* .*... *.*.. .*.* :  6
8         | *.*.. .*.*. ..*.* ..... .... :  6
9         | .*.*. ..*.* ...*. *.... .... :  6
10        | ....* *.... **... .**.. .... :  6
11        | **... .**.. ..... ..... .... :  4
12        | ..*.* ...*. *.... ..... .... :  4
13        | ...*. *...* .*... ..... .... :  4
14        | ***** *.... ..... ..... .... :  6
           ..................................
#         | 33333 33333 33333 33333 3333
```

See the Porta manual for further explanations. The strong validity table at the end is recording all the facet-vertex incidences in this polyhedron.

Similarly you can convert from an inequality to a generator representation of a polyhedron by first creating a file with the inequalities called **filename.ieq** and then using the command **traf filename.ieq** to get an output file called **filename.ieq.poi** that contains the generators. Here is a simple example in $\mathbb{R}^2$ that you can check by hand.

```
DIM = 2

VALID
0 0

INEQUALITIES_SECTION
x1 >= 0
x2 >= 0
-2x1 + x2 <= 1
x1 - 2x2 <= 1
END
```

```
DIM = 2

CONE_SECTION
(  1) 1 2
(  2) 2 1

CONV_SECTION
(  1) 0 0
(  2) 0 1
(  3) 1 0

END
```

Definition 2.9. The **characteristic cone** or **recession cone** of a polyhedron $P \subseteq \mathbb{R}^n$ is the polyhedral cone

$$\text{rec.cone}(P) := \{\mathbf{y} \in \mathbb{R}^n \,:\, \mathbf{x} + \mathbf{y} \in P \text{ for all } \mathbf{x} \in P\}.$$

Exercise 2.10. (**HW**) Prove the following facts about a polyhedron $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$.

(1) $\text{rec.cone}(P) = \{\mathbf{y} \,:\, A\mathbf{y} \leq \mathbf{0}\}$,
(2) $\mathbf{y}$ belongs to rec.cone(P) if and only if there exists an $\mathbf{x} \in P$ such that $\mathbf{x} + \lambda\mathbf{y} \in P$ for all $\lambda \geq 0$,
(3) $P + \text{rec.cone}(P) = P$,
(4) P is a polytope if and only if rec.cone$(P) = \{\mathbf{0}\}$,
(5) if $P = Q + C$ for a polytope Q and a polyhedral cone C, then $C = \text{rec.cone}(P)$.

Exercise 2.11. Using 2.10 (1) and Porta check that the recession cone of the permutahedron in Example 2.8 is $\{0\}$.

Definition 2.12. The **lineality space** of a polyhedron P is the linear subspace

$$\text{lin.space}(P) := \text{rec.cone}(P) \cap -\text{rec.cone}(P).$$

Check that if $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$ then lin.space$(P) = \{\mathbf{y} \,:\, A\mathbf{y} = \mathbf{0}\}$ and it is the largest subspace contained in P.

Definition 2.13. A polyhedron P is **pointed** if its lineality space is $\{\mathbf{0}\}$.

The last definition we want to make is that of the dimension of a polyhedron. To do this, we first need some basics from *affine linear algebra*.

Definition 2.14. (1) An **affine linear combination** of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbb{R}^n$ is the sum $\sum_{i=1}^{p} \lambda_i \mathbf{a}_i$ where $\lambda_i \in \mathbb{R}$ and $\sum_{i=1}^{p} \lambda_i = 1$.
(2) The **affine hull** of a set $S \subseteq \mathbb{R}^n$, denoted as aff.hull(S), is the set of all affine combinations of finitely many points of S.

Example 2.15. The affine hull of two points $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{R}^n$ is the line through the two points. The affine hull of a set $S \subseteq \mathbb{R}^n$ is the union of all lines through any two points in the convex hull of S. Also, the affine hull of $S \subseteq \mathbb{R}^n$ is the union of all lines through any two points in it. Check these statements if needed.

Check that for a matrix A and vector $\mathbf{b}$, the set $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ is closed under affine combinations and is hence its own affine hull. If $\mathbf{b} \neq \mathbf{0}$ we call $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ an **affine subspace**. If $\mathbf{x}_0$ is such that $A\mathbf{x}_0 = \mathbf{b}$, then $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ is the translate of the linear subspace $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ by $\mathbf{x}_0$.

Definition 2.16. The **dimension** of $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ is defined to be the dimension of the linear subspace $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$.

In fact, if S is any subset of $\mathbb{R}^n$ then there exists a matrix A and vector $\mathbf{b}$ such that $\text{aff.hull}(S) = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$. Thus every affine hull has a well-defined dimension.

Proposition 2.17. *If $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ and $A'\mathbf{x} = \mathbf{b}'$ is the subsystem of inequalities in $A\mathbf{x} \leq \mathbf{b}$ that hold at equality on P, then the affine hull of P equals $\{\mathbf{x} \in \mathbb{R}^n : A'\mathbf{x} = \mathbf{b}'\}$.*

Proof. If $\mathbf{p}_1, \ldots, \mathbf{p}_t \in P$ then $A'\mathbf{p}_i = \mathbf{b}'$ for all $i = 1, \ldots, t$ and therefore, if $\sum_{i=1}^t \lambda_i = 1$ then $A'(\sum_{i=1}^t \lambda_i \mathbf{p}_i) = \sum_{i=1}^t \lambda_i A'\mathbf{p}_i = (\sum_{i=1}^t \lambda_i)\mathbf{b}' = \mathbf{b}'$. This implies that $\text{aff.hull}(P) \subseteq \{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$.

To show the reverse inclusion, suppose $\mathbf{x}_0$ satisfies $A'\mathbf{x}_0 = \mathbf{b}'$. If $\mathbf{x}_0 \in P$ then $\mathbf{x}_0 \in \text{aff.hull}(P)$ since every set S is contained in its affine hull. If $\mathbf{x}_0 \notin P$, then select a point $\mathbf{x}_1 \in P$ such that it satisfies all the remaining inequalities in $A\mathbf{x} \leq \mathbf{b}$ with a strict inequality. (Why should such an $\mathbf{x}_1$ exist in P?) Then the line segement joining $\mathbf{x}_0$ and $\mathbf{x}_1$ contains at least one more point in $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ and hence the line through $\mathbf{x}_1$ and $\mathbf{x}_0$ is in the affine hull of P which implies that $\mathbf{x}_0$ is in the affine hull of P. $\square$

Definition 2.18. The **dimension** of a polyhedron is the dimension of its affine hull. The dimension of the empty set is taken to be -1.

Therefore, to calculate the dimension of a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$, we first determine the largest subsystem $A'x \leq \mathbf{b}'$ in $A\mathbf{x} \leq \mathbf{b}$ that holds at equality on P. Then the dimension of P is $n - \text{rank}(A')$. For instance, the dimension of the permutahedron from Example 2.8 is three since there is exactly one equation in the inequality representation produced by Porta.

Definition 2.19. The points $\mathbf{a}_1, \ldots, \mathbf{a}_p \in \mathbb{R}^n$ are said to be **affinely dependent** if the vectors $\begin{pmatrix} 1 \\ \mathbf{a}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mathbf{a}_p \end{pmatrix}$ in $\mathbb{R}^{n+1}$ are linearly dependent and affinely independent otherwise. Therefore, $\mathbf{a}_1, \ldots, \mathbf{a}_p$ are affinely dependent if and only if there exists scalars $\lambda_1, \ldots, \lambda_p$, not all zero, such that $\sum \lambda_i = 0$ and $\sum \lambda_i \mathbf{a}_i = \mathbf{0}$.

Proposition 2.20. *The dimension of the affine hull of k affinely independent points in $\mathbb{R}^n$ is $k - 1$.*

Proof. Suppose $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{R}^n$ are affinely independent. Then $\begin{pmatrix} 1 \\ \mathbf{a}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mathbf{a}_k \end{pmatrix}$ in $\mathbb{R}^{n+1}$ are linearly independent and span a k-dimensional linear space in $\mathbb{R}^{n+1}$. Call this space L. Note that $\mathbf{a}$ is an affine combination of $\mathbf{a}_1, \ldots, \mathbf{a}_k$ if and only if there exists scalars $\lambda_1, \ldots, \lambda_k$ such that $\mathbf{a} = \sum_{i=1}^k \lambda_i \mathbf{a}_i$, $\sum_{i=1}^k \lambda_i = 1$ which is if and only if $\mathbf{a}$ lies in $L \cap \{(x_0, \mathbf{x}) \in \mathbb{R}^{n+1} : x_0 = 1\}$. Therefore $\text{aff.hull}(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ has dimension $k - 1$. $\square$

Propositions 2.17 and 2.20 together provide the usual strategy to compute the dimension of a polyhedron $P \subseteq \mathbb{R}^n$; P has dimension at least k if P contains $k+1$ affinely independent points and dimension at most k if it satisfies $n - k$ linearly independent equations.

DEFINITION 2.21. A matrix $A = (a_{ij}) \in \mathbb{R}^{d\times d}$ is said to be **doubly stochastic** if it has the following properties:

$$\sum_{i=1}^{d} a_{ij} = 1 \;\forall\; j = 1,\ldots,d, \;\; \sum_{j=1}^{d} a_{ij} = 1 \;\forall\; i = 1,\ldots,d, \;\; a_{ij} \geq 0 \;\; \forall\; 1 \leq i,j \leq d.$$

EXERCISE 2.22. (**HW**) The **Birkhoff polytope**, $B(d)$, $d \in \mathbb{N}$, is described as follows. Let S_d be the set of all permutations of the set $\{1,\ldots,d\}$. For each element (permutation) $\sigma \in S_d$, let X^σ be the corresponding permutation matrix.

$$X^\sigma_{i,j} := \begin{cases} 1 \text{ if } \sigma(i) = j \\ 0 \text{ otherwise} \end{cases}$$

Each X^σ is a 0/1 matrix with exactly one 1 in each row and column. The Birkhoff polytope $B(d)$ is the convex hull of all permutation matrices of size $d \times d$. We could record X^σ also as a 0/1 vector in $\mathbb{R}^{d^2}$ and hence we may assume that $B(d) \subset \mathbb{R}^{d^2}$. Using Porta compute the inequality representation of $B(d)$ for a few small values of d and note the dimension in each case. Prove that the dimension of $B(d)$ is $(d-1)^2$.

Show that the Birkhoff polytope $B(d)$ is precisely the set of all doubly stochastic matrices of size $d \times d$. In particular, every doubly stochastic matrix is a convex combination of permutation matrices of size $d \times d$.

CHAPTER 3

Faces of a Polyhedron

Consider a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$. For a non-zero vector $\mathbf{c} \in \mathbb{R}^n$ and a $\delta \in \mathbb{R}$, let $H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{cx} = \delta\}$ and $H_{\leq} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{cx} \leq \delta\}$. We say that H is a **supporting hyperplane** of P if $P \subseteq H_{\leq}$ and $\delta = \max\{\mathbf{cx} : \mathbf{x} \in P\}$.

DEFINITION 3.1. A subset F of P is called a **face** of P if either $F = P$ or $F = P \cap H$ for some supporting hyperplane H of P. All faces of P except P itself are said to be **proper**. The faces of P can be partially ordered by set inclusion. The maximal proper faces of P in this ordering are called **facets**.

REMARK 3.2. In the combinatorial study of polyhedra it is usual to include the empty set as a face of a polyhedron. This makes the partially ordered set of all faces of a polyhedron into a *lattice*. In this course, we do not include the empty set as a face of a polyhedron.

The problem maximize$\{\mathbf{cx} : \mathbf{x} \in P\}$ is a **linear program**. If this maximum is finite, then the set of optimal solutions of the linear program is the face $F = P \cap H$ of P. In this case, δ is called the **optimal value** of the linear program maximize$\{\mathbf{cx} : \mathbf{x} \in P\}$. We write $F = \mathrm{face}_{\mathbf{c}}(P)$ to denote this. The improper face $P = \mathrm{face}_{\mathbf{0}}(P)$. Every face of P is of the form $\mathrm{face}_{\mathbf{c}}(P)$ for some $\mathbf{c} \in \mathbb{R}^n$.

The *dual* of the linear program

$$\text{maximize}\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\}$$

is the linear program

$$\text{minimize}\{\mathbf{yb} : \mathbf{y}A = \mathbf{c},\ \ \mathbf{y} \geq \mathbf{0}\}.$$

A linear program is **infeasible** if the underlying polyhedron is empty and **unbounded** if it has no finite optimal value.

EXERCISE 3.3. Prove that the linear program $\max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\}$ is unbounded if and only if there exists a $\mathbf{y}$ in the recession cone of $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ with $\mathbf{cy} > 0$.

COROLLARY 3.4. *Given a polyhedron P, the linear program* $\max\{\mathbf{cx} : \mathbf{x} \in P\}$ *is bounded if and only if $\mathbf{c}$ lies in the polar of the recession cone of P.*

Linear programming has a famous duality theorem which we now state without proof.

THEOREM 3.5. **Linear Programming Duality.** *If* $\max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\}$ *has a finite optimal value then*

$$\max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\} = \min\{\mathbf{yb} : \mathbf{y}A = \mathbf{c},\ \ \mathbf{y} \geq \mathbf{0}\}.$$

PROPOSITION 3.6. *Let $F \subseteq P$. Then F is a face of P if and only if $F \neq \emptyset$ and $F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$ for some subsystem $A'\mathbf{x} \leq \mathbf{b}'$ of $A\mathbf{x} \leq \mathbf{b}$.*

PROOF. Suppose F is a face of P. Then $F \neq \emptyset$ and $F = P \cap H$ for some supporting hyperplane $H = \{\mathbf{x} : \mathbf{cx} = \delta\}$ of P. Further, $\delta = \max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\} = \min\{\mathbf{yb} : \mathbf{y}A = \mathbf{c},\ \mathbf{y} \geq \mathbf{0}\}$ by linear programming duality. Let $\mathbf{y}^*$ be an optimal solution of the dual program and $A'\mathbf{x} \leq \mathbf{b}'$ be the inequalities in $A\mathbf{x} \leq \mathbf{b}$ indexed by all i such that $y_i^* > 0$. If $\mathbf{x} \in P$ then

$$\mathbf{x} \in F \Leftrightarrow \mathbf{cx} = \delta \Leftrightarrow \mathbf{y}_0 A\mathbf{x} = \mathbf{y}_0\mathbf{b} \Leftrightarrow \mathbf{y}_0(A\mathbf{x} - \mathbf{b}) = \mathbf{0} \Leftrightarrow A'\mathbf{x} = \mathbf{b}'.$$

Hence $F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$.

On the other hand, if $\emptyset \neq F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$ for some subsystem $A'\mathbf{x} \leq \mathbf{b}'$ of $A\mathbf{x} \leq \mathbf{b}$, then take $\mathbf{c} := \sum\{\mathbf{a}' : \mathbf{a}' \text{ is a row of } A'\}$ and $\delta := \sum b_i'$. Then $F = \mathrm{face}_{\mathbf{c}}(P)$ since if $\mathbf{x} \in F$, then $\mathbf{cx} = \delta$ while for all other $\mathbf{x} \in P$, $\mathbf{cx} < \delta$ since at least one inequality in $A'\mathbf{x} \leq \mathbf{b}'$ is satisfied with the strict inequality by such an $\mathbf{x}$. $\square$

COROLLARY 3.7. (1) *A polyhedron has only finitely many faces.*
(2) *Every face of a polyhedron is again a polyhedron.*
(3) *If G is contained in a face F of a polyhedron P, then G is a face of P if and only if G is a face of F.*

Among the faces of F, the most important ones are the minimal and maximal proper faces under set inclusion. We look at minimal faces quite carefully now.

PROPOSITION 3.8. *Let $F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$ be a face of P. Then F is a minimal face if and only if F is the affine space $\{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$.*

PROOF. Suppose $F = \{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$. Then by Proposition 3.6, F has no proper faces since all inequalities describing F are equalities. Therefore, by Corollary 3.7, F is a minimal face of P.

To prove the converse, suppose $F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$ is a minimal face of P. Let $S = \{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$. Clearly, $F \subseteq S$. We need to prove that $F = S$. Since F is a minimal face of P we have that if $\mathbf{x} \in F$ then $\mathbf{x}$ satisfies all inequalities in $A\mathbf{x} \leq \mathbf{b}$ but not in $A'\mathbf{x} \leq \mathbf{b}'$ with a strict inequality.

Suppose $\hat{\mathbf{x}} \in F$ and $\bar{\mathbf{x}} \in S \backslash F$. Then the line segment $[\hat{\mathbf{x}}, \bar{\mathbf{x}}] \subset S$. Since $\bar{\mathbf{x}} \in S \backslash F$, $\bar{\mathbf{x}} \notin P$. Therefore there exists at least one inequality $\mathbf{ax} \leq \beta$ in $A\mathbf{x} \leq \mathbf{b}$ but not in $A'\mathbf{x} \leq \mathbf{b}'$ such that $\mathbf{a}\bar{\mathbf{x}} > \beta > \mathbf{a}\hat{\mathbf{x}}$. Now let $\lambda := (\mathbf{a}\bar{\mathbf{x}} - \beta)/(\mathbf{a}\bar{\mathbf{x}} - \mathbf{a}\hat{\mathbf{x}})$. Then $0 < \lambda < 1$ and $\mathbf{a}(\lambda\hat{\mathbf{x}} + (1-\lambda)\bar{\mathbf{x}}) = \beta$. Let λ^* be the minimum such λ over inequalities violated by $\bar{\mathbf{x}}$. Define $\mathbf{x}^* := \lambda^*\hat{\mathbf{x}} + (1-\lambda^*)\bar{\mathbf{x}}$. Then $\mathbf{x}^* \in S \cap P = F$ and $\mathbf{x}^*$ satisfies at equality a subsystem of $A\mathbf{x} \leq \mathbf{b}$ that properly contains $A'\mathbf{x} \leq \mathbf{b}'$ which contradicts the minimality of F. $\square$

EXERCISE 3.9. (**HW**) Prove that each minimal face of a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ is a translate of the lineality space, $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$, of P.

The above exercise shows that each minimal face of a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ has dimension equal to $n - \mathrm{rank}(A)$. In particular, if P is pointed, then each minimal face of P is just a point. These points are called the **vertices** of P. Each vertex is determined by n linearly independent equations from $A\mathbf{x} = \mathbf{b}$.

We can now refine the Affine Minkowski Theorem for polyhedra as follows.

THEOREM 3.10. *Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ with p minimal faces and let $\mathbf{x}_i$ be a point from the minimal face F_i. Then $P = \mathrm{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_p) + \mathrm{rec.cone}(P)$.*

PROOF. Let $Q = \mathrm{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_p) + \mathrm{rec.cone}(P)$. Then $Q \subseteq P$ and by the Affine Weyl Theorem (Theorem 2.3), Q is a polyhedron. Suppose $\bar{\mathbf{x}} \in P \backslash Q$. Then since $\bar{\mathbf{x}} \notin Q$, there is some inequality in the inequality description of Q, say $\mathbf{ax} \leq \beta$ that is violated by $\bar{\mathbf{x}}$. Therefore, $\max\{\mathbf{ax} : \mathbf{x} \in Q\} = \beta < \mathbf{a}\bar{\mathbf{x}} \leq \max\{\mathbf{ax} : \mathbf{x} \in P\} =: \bar{\beta}$. We now consider two cases:

Case (1) $\bar{\beta} < \infty$: Let F_i be a minimal face of P contained in $\text{face}_{\mathbf{a}}(P)$. Then for all $\mathbf{x} \in F_i$, $\mathbf{ax} = \bar{\beta} > \beta$. If we could show that $F_i \subseteq Q$, then $\mathbf{ax} \leq \beta$ for all $\mathbf{x} \in F_i$ which would contradict the previous statement. So we proceed to show that $F_i \subseteq Q$. If $F_i = \{\mathbf{x}_i\}$ then $F_i \subseteq Q$. Else, there exists a point $\mathbf{y}_i (\neq \mathbf{x}_i) \in F_i$ and therefore, the line segment $[\mathbf{x}_i, \mathbf{y}_i] \subseteq P$. But since F_i is a minimal face of P, by Proposition 3.8, F_i is an affine space and so, in fact, the line through $\mathbf{x}_i$ and $\mathbf{y}_i$ is contained in F_i and hence in P. This implies that $A(\lambda \mathbf{y}_i + (1-\lambda)\mathbf{x}_i) = A(\mathbf{x}_i + \lambda(\mathbf{y}_i - \mathbf{x}_i)) \leq \mathbf{b}$ for all $\lambda \in \mathbb{R}$ which in turn implies that $A(\mathbf{y}_i - \mathbf{x}_i) = \mathbf{0}$. Therefore, $\mathbf{y}_i - \mathbf{x}_i \in \text{rec.cone}(P) = \text{rec.cone}(Q)$ and hence, $\mathbf{y}_i = \mathbf{x}_i + (\mathbf{y}_i - \mathbf{x}_i) \in Q$.

Case (2) $\max\{\mathbf{ax} : \mathbf{x} \in P\}$ is unbounded: In this case, by Exercise 3.3 there must exist a vector $\mathbf{y} \in \text{rec.cone}(P)$ such that $\mathbf{ay} > 0$. But both P and Q have the same recession cone and hence by the same exercise, the linear program $\max\{\mathbf{ax} : \mathbf{x} \in Q\}$ is also unbounded which contradicts that the maximum is $\beta < \mathbf{a}\bar{\mathbf{x}} < \infty$.

□

For completeness, we state some results about the facets of a polyhedron without proofs. Facets are just as important as the minimal faces of a polyhedron and programs like **Polymake** convert between minimal faces and facets of a polyhedron.

An inequality $\mathbf{ax} \leq \beta$ in $A\mathbf{x} \leq \mathbf{b}$ is called an implicit equality (in $A\mathbf{x} \leq \mathbf{b}$) if $\mathbf{ax} = \beta$ for all $\mathbf{x}$ such that $A\mathbf{x} \leq \mathbf{b}$. Let $A'\mathbf{x} \leq \mathbf{b}'$ be the set of implicit equalities in $A\mathbf{x} \leq \mathbf{b}$ and $A''\mathbf{x} \leq \mathbf{b}''$ be the rest. An inequality $\mathbf{ax} \leq \beta$ in $A\mathbf{x} \leq \mathbf{b}$ is **redundant** in $A\mathbf{x} \leq \mathbf{b}$ if it is implied by the remaining contraints in $A\mathbf{x} \leq \mathbf{b}$. An inequality system is **irredundant** if it has no redundant constraints.

THEOREM 3.11. [**Sch86**, Theorem 8.1] *If no inequality in $A''\mathbf{x} \leq \mathbf{b}''$ is redundant in $A\mathbf{x} \leq \mathbf{b}$ then there exists a bijection between the facets of $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ and the inequalities in $A''\mathbf{x} \leq \mathbf{b}''$ given by $F = \{\mathbf{x} \in P : \mathbf{ax} = \beta\}$ for any facet F of P and an inequality $\mathbf{ax} \leq \beta$ from $A''\mathbf{x} \leq \mathbf{b}''$.*

COROLLARY 3.12. (1) *Each proper face of P is the intersection of facets of P.*
(2) *P has no proper faces if and only if P is an affine space.*
(3) *The dimension of any facet of P is one less than the dimension of P.*
(4) *If P is full-dimensional and $A\mathbf{x} \leq \mathbf{b}$ is irredundant, then $A\mathbf{x} \leq \mathbf{b}$ is the unique minimal constraint representation of P, up to multiplication of inequalities by a positive scalar.*

We now use Exercise 2.22 to prove a classical combinatorial theorem about bipartite graphs. This is an example of how a geometric object like a polyhedron can imply theorems about a purely combinatorial object like an abstract graph.

DEFINITION 3.13. Given an undirected graph $G = (V, E)$, a **matching** in G is a collection of edges in E such that no two edges share a vertex. A matching is **perfect** if every vertex of G is incident to some edge in the matching.

COROLLARY 3.14. *Every regular bipartite graph has a perfect matching.*

PROOF. Let $G = (V, E)$ be a regular bipartite graph of degree $r \geq 1$, on the vertex set V consisting of the two *color classes* V_1 and V_2. Then $|E| = r|V_1| = r|V_2|$ which implies that $|V_1| = |V_2|$. Assume that $V_1 = \{1, \ldots, n\}$ and $V_2 = \{n+1, \ldots, 2n\}$. Also, note that every perfect matching of the bipartite graph G can be recorded by an $n \times n$ permutation matrix. Now define a matrix $A = (a_{ij})$ as

$$a_{ij} := \begin{cases} \frac{1}{r} \text{ if } \ \{i, n+j\} \in E \\ 0 \ \text{ otherwise} \end{cases}$$

Check that A is doubly stochastic. Therefore, by Exercise 2.22, there exists some permutation matrix B of size $n \times n$ such that if $b_{ij} = 1$ then $a_{ij} > 0$. This matrix B indexes a perfect matching in G. $\square$

CHAPTER 4

The Perfect Matching Polytope

Let $G = (V, E)$ be an undirected graph and $w : E \to \mathbb{R}$ a function that assigns weight $w(e)$ to edge $e \in E$. A typical example of a problem in **combinatorial optimization** is to find a maximum or minumum weight perfect matching in G. We will assume that $|V|$ is even throughout this lecture since otherwise G has no perfect matchings.

The **maximum weight perfect matching problem** in G can be modeled via polyhedra as follows. Define the **characteristic vector** $\chi^M \in \{0,1\}^E$ of a matching $M \subseteq E$ in G as follows:

$$\chi_e^M := \begin{cases} 1 & \text{if } e \in M \\ 0 & \text{otherwise} \end{cases}$$

Definition 4.1. The polytope $P(G) := \operatorname{conv}(\chi^M : M \text{ is a perfect matching in } G)$ in $\mathbb{R}^E$ is called the **perfect matching polytope** of G.

Then the maximum weight perfect matching problem in G is the linear program

$$\max \sum w(e)x_e \, : \, \mathbf{x} = (x_e) \in P(G)$$

which can be solved using algorithms for linear programming if an explicit inequality description of $P(G)$ is known. In this lecture, we prove a famous theorem due to Jack Edmonds from 1965 that provides a complete inequality description of $P(G)$. This work initiated the use of polyhedral methods for problems in combinatorial optimization.

Given a vector $\mathbf{x} \in \mathbb{R}^E$ and a set $S \subseteq E$, let $\mathbf{x}(S) := \sum_{e \in S} x_e$. Further, for $W \subseteq V$, let $\delta(W) := \{e \in E : |e \cap W| = 1\}$ be the set of edges in E that have precisely one vertex in W. In other words, $\delta(W)$ is the collection of edges in G that "leave" W. In particular, if $v \in V$, then $\delta(v)$ is the collection of edges that are incident to v.

Note that the characteristic vector of a perfect matching in G satisfies the following sets of inequalities in $\mathbb{R}^E$:

$$x_e \geq 0 \text{ for all } e \in E \tag{2}$$

$$\mathbf{x}(\delta(v)) = 1 \text{ for all } v \in V \tag{3}$$

$$\mathbf{x}(\delta(W)) \geq 1 \text{ for all } W \subseteq V, |W| \text{ odd} \tag{4}$$

As a warm-up we tackle the case when G is bipartite which follows as a corollary of Exercise 2.22.

Theorem 4.2. *If G is bipartite then* $P(G) = \{\mathbf{x} \in \mathbb{R}^E : \mathbf{x} \text{ satsifies (1) and (2)}\}$.

Proof. Let $Q(G) = \{\mathbf{x} \in \mathbb{R}^E : \mathbf{x} \text{ satsifies (1) and (2)}\}$. Since $P(G) \subseteq Q(G)$, we just need to argue that $Q(G) \subseteq P(G)$. Suppose $\mathbf{x} \in Q(G)$. Let V_1 and V_2 be the two classes of vertices in the bipartite graph G. Then since $\mathbf{x}(\delta(v)) = 1$ for all $v \in V$, we get that

$$|V_1| = \sum_{v \in V_1} \mathbf{x}(\delta(v)) = \sum_{v \in V_2} \mathbf{x}(\delta(v)) = |V_2|$$

where the middle equality follows from $\sum_{v\in V_1} \mathbf{x}(\delta(v)) = \mathbf{x}(E) = \sum_{v \in V_2} \mathbf{x}(\delta(v))$. Let $V_1 = \{1, \ldots, n\}$ and $V_2 = \{n+1, \ldots, 2n\}$. Define the matrix $A \in \mathbb{R}^{n \times n}$ by $a_{ij} := 0$ if there is no edge between vertex i and vertex $n+j$ and $a_{ij} := x_e$ otherwise. Then check that A is doubly stochastic and so is a convex combination of permutation matrices of size $n \times n$. Every permutation matrix in this combination has its support inside the support of A and hence corresponds to a perfect matching in G. Thus we have expressed $\mathbf{x}$ (sitting in the non-zero entries of A) as a convex combination of characteristic vectors of perfect matchings in G and so $\mathbf{x} \in P(G)$. $\square$

For a general graph $G = (V, E)$, the inequalities in (1) and (2) are not enough to cut out $P(G)$. To see this take the graph

$$G = (\{1, 2, 3, a, b, c\}, \{\{1, 2\}, \{2, 3\}, \{1, 3\}, \{a, b\}, \{b, c\}, \{a, c\}\})$$

and $\mathbf{x} = (1/2, 1/2, 1/2, 1/2, 1/2, 1/2) \in \mathbb{R}^E$. Check that $\mathbf{x}$ satisfies (1) and (2) but $P(G) = \emptyset$ for this example.

THEOREM 4.3. *For a graph G with an even number of vertices,*

$$P(G) = \{\mathbf{x} \in \mathbb{R}^E : \mathbf{x} \text{ satisfies (1), (2) and (3)}\}.$$

PROOF. Let $Q(G) := \{\mathbf{x} \in \mathbb{R}^E : \mathbf{x} \text{ satisfies (1), (2) and (3)}\}$. We need to show that $Q(G) \subseteq P(G)$. Suppose not and G is a smallest graph (with respect to $|V| + |E|$) such that $Q(G) \not\subseteq P(G)$. Let $\mathbf{x}$ be a vertex of $Q(G)$ that is not contained in $P(G)$. Then for each $e \in E$, we have the following properties:

(1) $x_e > 0$ — otherwise, $\mathbf{x}$ lies in $Q(G) \cap \{\mathbf{x} : x_f = 0\}$ for some edge $f \in E$ and we have $Q(G) \cap \{\mathbf{x} : x_f = 0\} \not\subseteq P(G) \cap \{\mathbf{x} : x_f = 0\}$. But these intersections correspond to $Q(G')$ and $P(G')$ for the smaller graph G' obtained from G by deleting the edge f, which contradicts the minimality of G.
(2) $x_e < 1$ — suppose there is an edge $f = \{v, w\} \in E$ such that $x_f = 1$. Then for all other edges e incident to v and w, $x_e = 0$ because $\mathbf{x}$ satisfies $\mathbf{x}(\delta(v)) = \mathbf{x}(\delta(w)) = 1$. Since for all $e \in E$, $0 \leq x_e \leq 1$ are valid inequalities for $Q(G)$, $\mathbf{x}$ lies on the face of $Q(G)$ at which $x_f = 1$ and $x_e = 0$ for all $e \neq f$ incident at v and w. This implies that $Q(G') \not\subseteq P(G')$ where G' is obtained from G by deleting all edges incident to v and w from E and v, w from V. Again we contradict the minimality of G.
(3) $|E| > |V|$ — else,
    (a) G is disconnected in which case the components of G give smaller counterexamples,
    (b) or G has a vertex of degree one in which case, the edge e incident to this vertex will be present in every perfect matching of G allowing us to delete the vertex and edge from G to get a smaller counterexample,
    (c) or G is an even cycle in which case the theorem is trivial.

Since $\mathbf{x}$ is a vertex of $Q(G)$ which lives in $\mathbb{R}^E$, there are $|E|$ linearly independent inequalities from (1)–(3) that hold at equality on $\mathbf{x}$. By (1), no inequality from (1) can hold at equality at $\mathbf{x}$, and since $|E| > |V|$, there must exist at least one inequality from (3) that holds at equality at $\mathbf{x}$. So there is a $W \subset V$, $|W|$ odd such that $\mathbf{x}(\delta(W)) = 1$. Note that for this W, we also need $|W| \geq 3$ since otherwise $|W| = 1$ and the equality $\mathbf{x}(\delta(W)) = 1$ is contained in (2). Similarly, $|V\backslash W| \geq 3$ since otherwise, $\mathbf{x}(\delta(V\backslash W)) = 1$ is contained in (2). These two (in)equalities are equivalent since $\delta(W) = \delta(V\backslash W)$ are the edges in the cut between W and $V\backslash W$.

Let G^1 and G^2 be obtained by contracting $V\backslash W$ and W respectively. Assume $V\backslash W$ is replaced by the vertex w^* in G^1 and W by w_* in G^2. For instance, G^1 has vertex set $W \cup \{w^*\}$ and edge set

$$\{e \,:\, e \subset W\} \cup \{\{v, w^*\} \,:\, v \in W \text{ and } \{v, w\} \in E \text{ for some } w \in V\backslash W\}.$$

Let $\mathbf{x}^1$ and $\mathbf{x}^2$ be the "projections" of $\mathbf{x}$ onto $\mathbb{R}^{E(G^1)}$ and $\mathbb{R}^{E(G^2)}$ respectively. By this we mean the following:

$$\mathbf{x}^1 \in \mathbb{R}^{E(G^1)},\ x^1_e := x_e \text{ if } e \subseteq W,\ x^1_{\{v,w^*\}} := \sum_{w \in V\backslash W} x_{\{v,w\}}.$$

Define $\mathbf{x}^2$ similarly.

Now we check that $\mathbf{x}^1 \in Q(G^1)$:

(1) By construction, $\mathbf{x}^1 \geq 0$ since $\mathbf{x} \geq 0$.
(2) For $v \in W$, $\mathbf{x}^1(\delta(v)) = \sum_{w \in W} x^1_{\{v,w\}} + x^1_{\{v,w^*\}} = \mathbf{x}(\delta(v)) = 1$.
Further, $\mathbf{x}^1(\delta(w^*)) = \sum_{v \in W} x^1_{\{v,w^*\}} = \mathbf{x}(\delta(W)) = 1$ because of our particular choice of W.
(3) Lastly, for any $U \subseteq W \cup \{w^*\}$, $|U|$ odd, if $w^* \notin U$, then $\mathbf{x}^1(\delta(U)) = \mathbf{x}(\delta(U)) \geq 1$. If $w^* \in U$, then $U = \{w^*\} \cup U'$ where $U' \subseteq W$ and $|U'|$ is even. Since $|W|$ is odd, we get that $|U'' := W\backslash U'|$ is odd. The set of edges that leave U is the union of the set of edges that go from U' to U'' which we will denote as T and the and the set of edges from w^* to U'' which we denote as S. Now T is the set of all edges that leave U'' minus the edges between U'' and $V\backslash W$ and S is the set of all edges that leave $V\backslash W$ minus the edges between $V\backslash W$ and U'. Therefore,

$$\mathbf{x}^1(\delta(U)) = \mathbf{x}(\delta(U'')) - \mathbf{x}(U'', V\backslash W) + \mathbf{x}(\delta(V\backslash W)) - \mathbf{x}(V\backslash W, U')$$

where the notation $\mathbf{x}(G, H)$ means the sum of all coordinates of $\mathbf{x}$ indexed by edges that go between the set of edges G and H. The two negative terms in the sum add to $-\mathbf{x}(\delta(V\backslash W))$ which cancels with the second positive term. Therefore, the right hand side is just $\mathbf{x}(\delta(U'')) \geq 1$ since $|U''|$ is odd.

Similarly, $\mathbf{x}^2 \in Q(G^2)$.

Since G^1 and G^2 are smaller graphs, we must have that $\mathbf{x}^1$ is a convex combination of characteristic vectors of perfect matchings in G^1. Similarly for $\mathbf{x}^2$. Suppose $\mathbf{x}^1 = \sum \lambda_i \chi^{M^1_i}$. Then there is a positive integer K that can be used to clear denominators such that $K\mathbf{x}^1 = \sum (K\lambda_i)\chi^{M^1_i}$ where $K\lambda_i$ are positive integers. This means that $K\mathbf{x}^1$ is a sum of $\sum K\lambda_i = K\sum \lambda_i = K$ characteristic vectors of perfect matchings in G^1, possibly with repetitions. Call these matchings $M^1_1, \ldots, M^1_K$. You could do the same thing to $\mathbf{x}^2$ with a positive integer K'. But note that we can also choose the same (large enough) positive integer to clear denominators in both cases. Call this common scalar K again. Then $K\mathbf{x}^2$ is the sum of characteristic vectors of perfect matchings $M^2_1, \ldots, M^2_K$ in G^2.

The last step is to use the above perfect matchings in G^1 and G^2 to write $\mathbf{x}$ as a convex combination of perfect matchings in G contradicting that $\mathbf{x} \notin P(G)$.

The perfect matching M^1_j in G_1 corresponds to a matching N^1_j in G that covers all vertices in W and exactly one vertex in $V\backslash W$. If f is an edge in G^1 then the number of

M_j^1's that contain f equals Kx_f^1. If $f \subseteq W$, then $Kx_f^1 = Kx_f$. Else, $f = \{v, w^*\}$ for some $v \in W$ and $Kx_f^1 = K\sum_{w \in V \setminus W, \{v,w\} \in E} x_{\{v,w\}} = \sum_{w \in V \setminus W, \{v,w\} \in E} Kx_{\{v,w\}}$. Thus we may assume that for any edge $e \in E$ with $|e \cap W| \geq 1$, e is contained in Kx_e of the N_j^1's. (Note that we can choose K large enough that Kx_e is an integer.) Similarly, each perfect matching M_j^2 in G^2 corresponds to a matching N_j^2 in G that covers all vertices in $V \setminus W$ and exactly one vertex in W and we may assume that each $e \in E$ with $|e \cap V \setminus W| \geq 1$, appears in precisely Kx_e of the N_j^2's.

Now take an edge $e \in E$ that lies in the cut $\delta(W)$. This edge appears in precisely Kx_e of the N_j^1's and Kx_e of the N_j^2's. These particular matchings do not contain any other edges in $\delta(W)$. Further, every N_j^i contains exactly one edge from $\delta(W)$. Hence we may assume that for each j, N_j^1 and N_j^2 contain the same unique edge in $\delta(W)$. Then $N_j^1 \cup N_j^2$ is a perfect matching in G. Further, each $e \in E$ is in exactly Kx_e of these perfect matchings $N_j^1 \cup N_j^2$. Therefore, $K\mathbf{x}$ is the sum of the characteristic vectors of these K perfect matchings $N_j^1 \cup N_j^2$ (each used exactly once) which makes $\mathbf{x}$ a convex combination of characteristic vectors of perfect matchings in G. $\square$

Exercise 4.4. (**HW**) Prove why we can assume $|E| > |V|$ in the proof of Theorem 4.3.

As a corollary to Theorem 4.3, one can derive Edmond's inequality description of the **matching polytope** of an undirected graph $G = (V, E)$. For $W \subseteq V$, let $E(W)$ denote the edges of G that involve pairs of vertices in W. The matching polytope of G, denoted as $M(G)$, is the convex hull of all characteristic vectors of matchings in G. It is a $0-1$-polytope in $\mathbb{R}^E$. Check that the characteristic vector of a matching in G satisfies the following inequalities:

$$x_e \geq 0 \quad \forall \quad e \in E \tag{5}$$

$$\mathbf{x}(\delta(v)) \leq 1 \quad \forall \quad v \in V \tag{6}$$

$$\mathbf{x}(E(W)) \leq \frac{1}{2}(|W| - 1) \quad \forall \quad W \subseteq V, |W| \text{ odd} \tag{7}$$

Theorem 4.5. *The matching polytope of $G = (V, E)$ is*

$$M(G) = \left\{\mathbf{x} \in \mathbb{R}^E : \mathbf{x} \text{ satisfies } (5), (6), (7)\right\}.$$

Exercise 4.6. (**HW**) Prove Theorem 4.5. One possible strategy is the following: Suppose $\mathbf{x} \in \mathbb{R}^E$ satisfies (5),(6) and (7). Let $G^* = (V^*, E^*)$ be a disjoint copy of G where the copy of vertex v is v^* and the copy of $e = \{v, w\}$ is $e^* = \{v^*, w^*\}$. Let $\tilde{G}$ be the graph with vertex set $V \cup V^*$ and edge set $E \cup E^* \cup \left\{\{v, v^*\} : v \in V\right\}$. Define $\tilde{\mathbf{x}}(e) := \tilde{\mathbf{x}}(e^*) := \mathbf{x}(e)$ for $e \in E$ and $\tilde{\mathbf{x}}(\{v, v^*\}) := 1 - \mathbf{x}(\delta(v))$ for $v \in V$.

- Verify that $\tilde{\mathbf{x}}$ lies in the perfect matching polytope of $\tilde{G}$.
- Thus $\tilde{\mathbf{x}}$ is a convex combination of characteristic vectors of perfect matchings in $\tilde{G}$. By restricting to $\mathbf{x}$ and G, show that $\mathbf{x}$ is in $M(G)$.

CHAPTER 5

Elementary Complexity Theory

In this lecture we give a very informal introduction to the theory of computational complexity so that we can analyze the objects and algorithms we will see in this course. The treatment is more intuitive than precise.

The basic goal of complexity theory is to create a measure of the complexity or difficulty of solving a problem. As a first step, we need to encode the problem using an alphabet.

DEFINITION 5.1. (1) The *alphabet* Σ is a finite set.
(2) Elements of Σ are called *letters* or *symbols*.
(3) A *string* or *word* from Σ is an ordered finite sequence of letters from Σ. The empty word is denoted as $\emptyset$. The set of all words from Σ is denoted as Σ^*.
(4) The *size* of a word is the number of letters in it. The empty word has size zero.

The objects we are interested in are usually numbers, vectors, matrices etc. These objects need to be encoded as words in Σ and hence will have a size.

EXAMPLE 5.2. In the binary encoding of numbers, the alphabet is $\Sigma = \{0, 1\}$ and a positive number p is expressed in base two and has size equal to $\lceil \log_2(|p| + 1)\rceil$. For example, 32 has the binary encoding 100000 which has size $6 = \lceil \log_2(33)\rceil$ while 31 has encoding 11111 which has size $5 = \lceil \log_2(32)\rceil$.

Different alphabets and encoding schemes express the same object via different strings and each such string has a size in that encoding. There are transformations between encoding schemes which also transform between sizes. Under reasonable assumptions all these sizes are linearly equivalent, or more formally, $\text{size}_1 = \mathcal{O}(\text{size}_2)$. For our purposes we fix a concept of size of a string inspired by the binary encoding of numbers.

(1) For an integer p, $\text{size}(p) := 1 + \lceil \log_2(|p| + 1)\rceil$ where the first bit encodes the sign of the number and the rest encodes the absolute value of the number.
(2) For a rational number $\alpha = \frac{p}{q}$, we may assume that p and q are relatively prime integers and $q \geq 1$. Then $\text{size}(\alpha) := 1 + \lceil \log_2(|p| + 1)\rceil + \lceil \log_2(|q| + 1)\rceil$.
(3) For a rational vector $\mathbf{c} = (c_1, \ldots, c_n)$, $\text{size}(\mathbf{c}) := n + \sum_{i=1}^{n} \text{size}(c_i)$.
(4) For a rational matrix $A = (a_{ij}) \in \mathbb{Q}^{m\times n}$, $\text{size}(A) := mn + \sum_{i,j} \text{size}(a_{ij})$.
(5) The size of a linear inequality $\mathbf{ax} \leq \beta$ or equation $\mathbf{ax} = \beta$ is $1 + \text{size}(\mathbf{a}) + \text{size}(\beta)$.
(6) The size of $A\mathbf{x} \leq \mathbf{b}$ or $A\mathbf{x} = \mathbf{b}$ is $1 + \text{size}(A) + \text{size}(\mathbf{b})$.

A *problem* is a question or a task. For example, the problem could be *Does the system $A\mathbf{x} \leq \mathbf{b}$ have a solution?* or *Find a solution of the system $A\mathbf{x} \leq \mathbf{b}$ or determine that there is none.* The former is called a *decision problem* since the answer is a "yes" or "no". Formally, we think of a problem as a set $\Pi \subseteq \Sigma^* \times \Sigma^*$ and the job is, given a string $z \in \Sigma^*$, find a string $y \in \Sigma^*$ such that $(z, y) \in \Pi$ or decide that no such y exists. The string z is an *input* or *instance* of the problem while y is an *output* or *solution* to the problem. A decision problem can then be recorded as the set of tuples $\{(z, \emptyset)\}$ as z ranges over all instances of the problem for which the answer to the problem is "yes". In other words, $(z, \emptyset) \notin \Pi$ if and only if z is an instance of the problem for which the answer to the problem is "no". Going back to our examples, the problem *Does the system* $A\mathbf{x} \leq \mathbf{b}$ *have a solution?* is the set $\Pi = \{((A, \mathbf{b}), \emptyset) \,:\, A\mathbf{x} \leq \mathbf{b} \text{ is feasible}\}$, and the problem *Find a solution of the system* $A\mathbf{x} \leq \mathbf{b}$ *or determine that there is none* is the set $\Pi' = \{((A, \mathbf{b}), \mathbf{x}) \,:\, \mathbf{x} \text{ satisfies } A\mathbf{x} \leq \mathbf{b}\}$. Therefore, if $A\mathbf{x} \leq \mathbf{b}$ is infeasible then the tuple $((A, \mathbf{b}), \emptyset)$ would not belong to Π and $((A, \mathbf{b}), *)$ would not belong to Π' where $*$ means "anything".

An *algorithm* A for a problem Π is a list of instructions that will "solve" Π. By this we mean that given a string $z \in \Sigma^*$, A will either find a solution y of the problem Π (i.e., find a y such that $(z, y) \in \Pi$) or stop without an output if no such y exists. We are interested in the *running times* of algorithms. Since we don't want this to depend on the particular implementation of the algorithm or the speed of the computer being used, we need a definition of running time that is intrinsic to the problem and algorithm. The way out is to define the running time of an algorithm on an input z to be the number of *elementary bit operations* needed by the algorithm before it stops, given the input z.

Definition 5.3. The *running time function* of an algorithm A for a problem Π is the function $f_A : \mathbb{N} \to \mathbb{N}$ such that

$$f(s) := \max_{\{z \,:\, \text{size}(z) \leq s\}}(\text{running time of } A \text{ for input } z).$$

Note that we can always assume that running time functions are monotonically increasing.

Definition 5.4. An algorithm is said to be *polynomial-time or polynomial* if its running time function $f_A(s)$ is bounded above by a polynomial in s. A problem Π is *polynomial-time solvable* if it has an algorithm that is polynomial.

The elementary arithmetic operations are *addition, subtraction, multiplication, division* and *comparison of numbers*. In rational arithmetic these can be executed by polynomial-time algorithms. Therefore, to decide if an algorithm is polynomial-time, it is enough to show that the number of elementary operations needed by the algorithm is bounded by a polynomial in the size of the input and that the sizes of all intermediate numbers created are also polynomially bounded in the size of the input.

We now very informally describe the problem classes we are interested in.

(1) The class of all decision problems that are polynomial-time solvable is denoted as $\mathcal{P}$.
(2) The class of all decision problems for which a string y can be verified to be a solution for an instance z in polynomial time is called $\mathcal{NP}$. In particular, the size of the "guess" y has to be polynomially bounded in the size of z. It is not important how y is produced.
(3) The class of decision problems for which a string z can be verified to not be an instance of the problem in polynomial time is called $co - \mathcal{NP}$.

Therefore, $\mathcal{NP} \cap co - \mathcal{NP}$ consists of those decision problems for which a positive or negative answer to a given instance z can be verfied in polynomial-time in the size of z. If a decision problem is in $\mathcal{P}$ then it is in $\mathcal{NP} \cap co - \mathcal{NP}$ since we are sure to either have a solution or decide there is none in polynomial time. However, it is a big open problem (worth a million U.S. dollars) whether $\mathcal{P} = \mathcal{NP}$, or even whether $\mathcal{NP} = co - \mathcal{NP}$.

The $\mathcal{NP}$*-complete* problems are the hardest problems among all $\mathcal{NP}$ problems in the sense that all problems in $\mathcal{NP}$ can be "reduced" to an $\mathcal{NP}$-complete problem. By "reducibile" we mean that there is a polynomial-time algorithm that will convert instances of one problem into instances of another. So if an $\mathcal{NP}$-complete problem had a polynomial-time algorithm then $\mathcal{P} = \mathcal{NP}$. The prototypical example of an $\mathcal{NP}$-complete problem is the *integer linear program*

$$\max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \text{ integer}\}.$$

To finish this lecture, we look at some examples of problems that are in $\mathcal{NP} \cap co - \mathcal{NP}$ and, in fact, in $\mathcal{P}$.

Example 5.5. Consider the fundamental problem of linear algebra:
Π_1: *given a rational matrix A and a rational vector* $\mathbf{b}$ *does* $A\mathbf{x} = \mathbf{b}$ *have a solution?*

(1) We first prove that if a rational linear system $A\mathbf{x} = \mathbf{b}$ has a solution, then it has one of size polynomially bounded by the size of A and $\mathbf{b}$. This requires several steps, some of which we state without proofs and some as exercises.

    Exercise 5.6. (**HW**) Let A be a square rational matrix of size s. Show that the size of $\det(A)$ is at most $2s$. (On the other hand, $\det(A)$ itself can be exponential in the size of A. Find such an example.)

    Corollary 5.7. *The inverse A^{-1} of a non-singular square rational matrix A has size polynomially bounded by the size of A.*

    Proof. The entries of A^{-1} are quotients of subdeterminants of A. $\square$

    Theorem 5.8. *If $A\mathbf{x} = \mathbf{b}$ has a solution, it has one of size polynomially bounded by the size of A and $\mathbf{b}$.*

    Proof. Assume that the rows of A are linearly independent and that $A = [A_1\ A_2]$ with A_1 non-singular. Then $(A_1^{-1}\mathbf{b}, \mathbf{0})$ is a solution of $A\mathbf{x} = \mathbf{b}$ of the size needed. $\square$

(2) Now it is easy to see that Π_1 is in $\mathcal{NP}$ since by (1), a solution $\mathbf{x}$ of polynomial size exists and clearly, we can plug this $\mathbf{x}$ into $A\mathbf{x} = \mathbf{b}$ to check that it is indeed a solution, in polynomial time.
(3) By Exercise 1.12, either $A\mathbf{x} = \mathbf{b}$ has a solution or there exists a $\mathbf{y}$ such that $\mathbf{y}A = \mathbf{0}$ but $\mathbf{yb} \neq 0$. By Theorem 5.8, the system $\mathbf{y}A = \mathbf{0}$, $\mathbf{yb} = 1$ has a solution of size polynomially bounded by the sizes of A and $\mathbf{b}$. Such a solution can be verified to be a solution in polynomial time. Thus $\Pi_1 \in co - \mathcal{NP}$.
(4) It turns out that Π_1 is actually in $\mathcal{P}$ which subsumes the above results. This is done by proving that *Guassian elimination* is a polynomial-time algorithm for solving linear equations.

Example 5.9. Consider the problem
Π_2*: given a rational matrix A and a rational vector* $\mathbf{b}$ *does* $A\mathbf{x} \leq \mathbf{b}$ *have a solution?*
As for Π_1, we prove that $\Pi_2 \in \mathcal{NP} \cap co - \mathcal{NP}$. In fact, $\Pi_2 \in \mathcal{P}$.

Exercise 5.10. (**HW**) Prove that if the rational system $A\mathbf{x} \leq \mathbf{b}$ has a solution it has one of size polynomially bounded by the size of A and $\mathbf{b}$.

(**Hint**: Use the fact that the polyhedron $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ has a minimal face $F = \{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$ for a subsytem $A'\mathbf{x} \leq \mathbf{b}'$ of $A\mathbf{x} \leq \mathbf{b}$ (Proposition 3.8) and Theorem 5.8.)

This proves that Π_2 is in $\mathcal{NP}$.

EXERCISE 5.11. (**HW**) Use the variant of Farkas Lemma in Exercise 1.13 (2) to prove that $\Pi_2 \in co-\mathcal{NP}$.

The fact that Π_2 is in $\mathcal{P}$ is far more sophisticated than the corresponding proof for Π_1. It follows from the fact that linear programming is in $\mathcal{P}$, a problem that was open for about forty years until Kachiyan found the *ellipsoid algorithm* for solving linear programs in the early 1980's.

CHAPTER 6

Complexity of Rational Polyhedra

Lemma 6.1. [**Sch86**, Corollary 3.2d] *Let $A \in \mathbb{Q}^{m\times n}$ and $\mathbf{b} \in \mathbb{Q}^m$ such that each row of the matrix $[A\ \mathbf{b}]$ has size at most ϕ. If $A\mathbf{x} = \mathbf{b}$ has a solution, then*

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} = \{\mathbf{x}_0 + \lambda_1\mathbf{x}_1 + \cdots + \lambda_t\mathbf{x}_t : \lambda_1, \ldots, \lambda_t \in \mathbb{R}\}$$

for certain vectors $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_t$ of size at most $4n^2\phi$.

Proof. The equation on display in the lemma is writing the affine space $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ as the translate of the linear space $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ by a vector $\mathbf{x}_0$ in the affine space. Therefore, $A\mathbf{x}_0 = \mathbf{b}$ and $A\mathbf{x}_i = \mathbf{0}$ for $i = 1, \ldots, t$ with $\mathbf{x}_1, \ldots, \mathbf{x}_t$ a basis for $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$.

We may assume that the rows of A are linearly independent and hence $A = [A_1\, A_2]$ where A_1 is non-singular. Also, then $t = n - m$. We can take $\mathbf{x}_0 := (A_1^{-1}\mathbf{b}, \mathbf{0})$ and $\mathbf{x}_1, \ldots, \mathbf{x}_t$ as the columns of the matrix $\begin{pmatrix} A_1^{-1}A_2 \\ -I \end{pmatrix}$. Then the non-zero components of the $\mathbf{x}_i$'s are all quotients of subdeterminants of the matrix $[A\,\mathbf{b}]$ (by Cramer's rule) of order at most $m \le n$. The size of $[A\ \mathbf{b}]$ is at most $m + m\phi \le n\phi$. Then by Exercise 5.6, the size of every subdeterminant of $[A\ \mathbf{b}]$ is at most $2n\phi$. Therefore, each component of $\mathbf{x}_i$, $i = 0, \ldots, t$ has size at most $4n\phi$ and hence, each $\mathbf{x}_i$ has size at most $4n^2\phi$. □

Let C be a polyhedral cone. The only minimal face of C is its lineality space. Let t be the dimension of lin.space(C). A face of C of dimension $t+1$ is called a **minimal proper face** of C. So if C is pointed, then $t = 0$ and the minimal proper faces of C are its *extreme rays*.

Exercise 6.2. (**HW**) Let $C = \{\mathbf{x} : A\mathbf{x} \le \mathbf{0}\}$. Prove that if G is a minimal proper face of C then $G = \{\mathbf{x} : A'\mathbf{x} = \mathbf{0}, \mathbf{a}\mathbf{x} \le 0\}$ where $A'\mathbf{x} \le \mathbf{0}$ is a subsystem of $A\mathbf{x} \le \mathbf{0}$ and $\mathbf{a}$ is a row of A such that rank $\begin{pmatrix} A' \\ \mathbf{a} \end{pmatrix} = n - \dim(\text{lin.space}(C))$ and lin.space$(C) = \{\mathbf{x} : \mathbf{a}\mathbf{x} = 0, A'\mathbf{x} = \mathbf{0}\}$.

Exercise 6.3. For a polyhedral cone C with minimal proper faces $G_1, \ldots, G_s$, choose for each $i = 1, \ldots, s$ a vector $\mathbf{y}_i \in G_i \backslash \text{lin.space}(C)$ and vectors $\mathbf{z}_0, \ldots, \mathbf{z}_t$ in lin.space(C) such that lin.space$(C) = \text{cone}(\mathbf{z}_0, \ldots, \mathbf{z}_t)$. Then prove that

$$C = \text{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s, \mathbf{z}_0, \ldots, \mathbf{z}_t).$$

Theorem 6.4. [**Sch86**, Theorem 8.5] *Let $P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ be a non-empty polyhedron.*

(1) *For each minimal face F of P, choose a vector $\mathbf{x}_F \in F$;*
(2) *For each minimal proper face F of* rec.cone(P) *choose a vector $\mathbf{y}_F \in F \backslash \text{lin.space}(P)$;*
(3) *Choose a generating set $\mathbf{z}_1, \ldots, \mathbf{z}_t$ of* lin.space(P). *Then,*

$$P = \text{conv}(\mathbf{x}_F \text{ from (1)}) + \text{cone}(\mathbf{y}_F \text{ from (2)}) + \text{lin.span}(\mathbf{z}_1, \ldots, \mathbf{z}_t \text{ from (3)}).$$

EXERCISE 6.5. (**HW**) Let $P = \text{conv}(\mathbf{b}_1, \ldots, \mathbf{b}_p) + \text{cone}(\mathbf{c}_1, \ldots, \mathbf{c}_q) \subseteq \mathbb{R}^n$ be a full-dimensional polyhedron. Prove that each facet of P is determined by a linear equation of the form

$$\det \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \mathbf{x} & \mathbf{b}_{i_1} & \cdots & \mathbf{b}_{i_k} & \mathbf{c}_{j_1} & \cdots & \mathbf{c}_{j_{n-k}} \end{bmatrix} = \mathbf{0} \tag{8}$$

where $\mathbf{x}$ is the vector of variables $\mathbf{x} = (x_1, \ldots, x_n)$.

DEFINITION 6.6. Let $P \subseteq \mathbb{R}^n$ be a rational polyhedron.

(1) The **facet complexity** of P is the smallest number ϕ such that $\phi \geq n$ and there exists a system $A\mathbf{x} \leq \mathbf{b}$ describing P where each inequality has size at most ϕ.
(2) The **vertex complexity** of P is the smallest number ν such that $\nu \geq n$ and there exists rational vectors $\mathbf{b}_1, \ldots, \mathbf{b}_p$ and $\mathbf{c}_1, \ldots, \mathbf{c}_q$ each of size at most ν such that $P = \text{cone}(\mathbf{b}_1, \ldots, \mathbf{b}_p) + \text{conv}(\mathbf{c}_1, \ldots, \mathbf{c}_q)$.

THEOREM 6.7. [**Sch86**, Theorem 10.2] *Let $P \subseteq \mathbb{R}^n$ be a polyhedron with facet complexity ϕ and vertex complexity ν. Then $\nu \leq 4n^2\phi$ and $\phi \leq 4n^2\nu$.*

PROOF. We first prove that $\nu \leq 4n^2\phi$. Since the facet complexity of P is ϕ, $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ with the size of each inequality at most ϕ. Note the following facts.

(1) By Proposition 3.8, each minimal face of P is of the form $\{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$ for some subsystem $A'\mathbf{x} \leq \mathbf{b}'$ of $A\mathbf{x} \leq \mathbf{b}$. By Lemma 6.1, each minimal face contains a vector of size at most $4n^2\phi$.
(2) The lineality space of P is $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$. Again, by Lemma 6.1, this linear space is generated by vectors of size at most $4n^2\phi$.
(3) Let F be a minimal proper face of rec.cone(P). Then by Exercise 6.2, F is of the form $\{\mathbf{x} : A'\mathbf{x} = \mathbf{0}, \mathbf{a}\mathbf{x} \leq 0\}$. Choose $\mathbf{y}_F$ such that $A'\mathbf{y}_F = \mathbf{0}$ and $\mathbf{a}\mathbf{y}_F = -1$ say. Then $\mathbf{y}_F$ is not in the lineality space of P and has size at most $4n^2\phi$.

Now use Theorem 6.4 to conclude that the facet complexity of P is at most $4n^2\phi$.

Next we argue that $\phi \leq 4n^2\nu$. If $n = 1$ then $P \subseteq \mathbb{R}$ and we need at most two inequalities to describe P of the form $x \leq \alpha$ where α is a vertex of P. Each of these inequalities has size at most $1 + \nu$ and hence the result holds. So assume that $n \geq 2$. Suppose $P = \text{conv}(\mathbf{b}_1, \ldots, \mathbf{b}_p) + \text{cone}(\mathbf{c}_1, \ldots, \mathbf{c}_q)$ where the $\mathbf{b}_i$'s and $\mathbf{c}_j$'s are rational vectors of size at most ν.

Suppose P is full-dimensional. Then by Exercise 6.5 each facet of P is determined by a linear equation of the form

$$\det \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \mathbf{x} & \mathbf{b}_{i_1} & \cdots & \mathbf{b}_{i_k} & \mathbf{c}_{j_1} & \cdots & \mathbf{c}_{j_{n-k}} \end{bmatrix} = \mathbf{0}. \tag{9}$$

Expanding this determinant by its first column we obtain

$$\sum_{i=1}^{n} (-1)^i (\det(D_i)) x_i = -\det(D_0)$$

where each D_i is an $n \times n$ submatrix of the matrix in (9). Each D_i has size at most $2n(\nu+1)$. Therefore, the equation and the corresponding inequality for the facet have size at most $4n^2\nu$.

If P is not full-dimensional, then as above, find equations of size at most $4n^2\nu$ defining the affine hull of P. (How do we do that?) Further, there exists $n - \dim(P)$ coordinates that can be deleted to make P full-dimensional. This projected polyhedron Q also has vertex complexity at most ν and hence can be described by linear inequalities of size at most $4(n - \dim(P))^2\nu$. By adding zero coordinates we can extend these inequalities to inequalities valid for P. Now add in the equations of the affine hull of P to get an inequality description of P of the required size. $\square$

CHAPTER 7

The Integer Hull of a Rational Polyhedron

Let $P \subseteq \mathbb{R}^n$ be a rational polyhedron. Recall that this means that we can assume

$$P = \{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} \leq \mathbf{b}\}$$

for some rational $m \times n$ matrix A and rational vector $\mathbf{b}$. By clearing denominators in $A\mathbf{x} \leq \mathbf{b}$ we may assume without loss of generality that $A \in \mathbb{Z}^{m \times n}$ and $\mathbf{b} \in \mathbb{Z}^m$.

Definition 7.1. If $P \subseteq \mathbb{R}^n$ is a rational polyhedron, then its **integer hull**

$$P^I := \operatorname{conv}(P \cap \mathbb{Z}^n)$$

is the convex hull of all integer vectors in P.

Theorem 7.2. *For any rational polyhedron $P \subseteq \mathbb{R}^n$, its integer hull P^I is again a polyhedron. If P^I is non-empty, then both P and P^I have the same recession cone.*

Note that the theorem is true if P is a polytope since $P \cap \mathbb{Z}^n$ is finite and hence its convex hull is a polytope by definition. Also if C is a rational cone then $C^I = C$ since C is generated by integer vectors.

Proof. Let $P = Q + C$ where Q is a polytope and C is the recession cone of P. Let $\mathbf{y}_1, \ldots, \mathbf{y}_s \in \mathbb{Z}^n$ generate C as a cone and consider the *parallelepiped/zonotope* (a polytope):

$$Z := \{\sum_{i=1}^{s} \mu_i \mathbf{y}_i \,:\, 0 \leq \mu_i \leq 1,\ i = 1, \ldots, s\}.$$

To prove the theorem we will show that $P^I = (Q+Z)^I + C$. Since $Q+Z$ is a polytope, so is $(Q+Z)^I$ and hence P^I will be a polyhedron. Also, if $P^I \neq \emptyset$, then C will be the recession cone of P^I.

- ($P^I \subseteq (Q+Z)^I + C$): Let $\mathbf{p} \in P \cap \mathbb{Z}^n$. Then $\mathbf{p} = \mathbf{q} + \mathbf{c}$ for some $\mathbf{q} \in Q$ and $\mathbf{c} \in C$. Also $\mathbf{c} = \sum \mu_i \mathbf{y}_i$ for some $\mu_i \geq 0$. Since $\mu_i = (\mu_i - \lfloor \mu_i \rfloor) + \lfloor \mu_i \rfloor$, we get that $\mathbf{c} = \sum (\mu_i - \lfloor \mu_i \rfloor)\mathbf{y}_i + \sum \lfloor \mu_i \rfloor \mathbf{y}_i = \mathbf{z} + \mathbf{c}'$ where $\mathbf{z} \in Z$ and $\mathbf{c}' \in C \cap \mathbb{Z}^n$. Therefore, $\mathbf{p} = (\mathbf{q} + \mathbf{z}) + \mathbf{c}'$ and hence $\mathbf{q} + \mathbf{z} = \mathbf{p} - \mathbf{c}' \in \mathbb{Z}^n$. This implies that $\mathbf{p} \in (Q+Z)^I + C$.
- ($P^I \supseteq (Q+Z)^I + C$): $(Q+Z)^I + C \subseteq P^I + C = P^I + C^I \subseteq (P+C)^I = P^I$. $\square$

Now that we have a fundamental grip on P^I, we can ask many questions such as these.

Problem 7.3.
1. How can we compute P^I given P? If $P^I = \{\mathbf{x} \in \mathbb{R}^n \,:\, M\mathbf{x} \leq \mathbf{d}\}$ then what is the dependence of M and $\mathbf{d}$ on A and $\mathbf{b}$?
2. What is the complexity of P^I?
3. How can we decide whether $P^I = \emptyset$? Is there a Farkas Lemma that can certify the existence or non-existence of an integer point in P?

(4) When does $P = P^I$?

We will see answers to these questions in later lectures. At this point, we should see an example of an integer hull of a rational polyhedron. However, since we do not yet know a systematic way to compute integer hulls, we will just have to do some ad hoc computations on small examples. Of course, if we know that $P = P^I$ then no computation is needed to calculate P^I.

DEFINITION 7.4. A rational polyhedron $P \subseteq \mathbb{R}^n$ is an **integral polyhedron** if $P = P^I$.

LEMMA 7.5. *For a rational polyhedron P, the following are equivalent:*

(1) $P = P^I$*;*
(2) *each face of P contains integral vectors;*
(3) *each minimal face of P contains integral vectors;*
(4) $\max\{\mathbf{cx} : \mathbf{x} \in P\}$ *is attained by an integral $\mathbf{x}$ for each $\mathbf{c}$ for which the max is finite;*
(5) $\max\{\mathbf{cx} : \mathbf{x} \in P\}$ *is an integer for each $\mathbf{c} \in \mathbb{Z}^n$ for which the max is finite;*
(6) *each rational supporting hyperplane of P contains an integral point.*

PROOF.

- (1) $\Rightarrow$ (2): Let $F = \{\mathbf{x} \in P : \mathbf{ax} = \beta\}$ be a face of P with $\mathbf{ax} < \beta$ for all $\mathbf{x} \in P \backslash F$. If $\bar{\mathbf{x}} \in F \subseteq P = P^I$, then $\bar{\mathbf{x}} = \sum \lambda_i \mathbf{x}^i$ with $\mathbf{x}^i \in P \cap \mathbb{Z}^n$, $\lambda_i \geq 0$ and $\sum \lambda_i = 1$. Thus $\beta = \mathbf{a} \cdot \bar{\mathbf{x}} = \sum \lambda_i \mathbf{a} \cdot \mathbf{x}^i \leq \beta$. This implies that $\mathbf{a} \cdot \mathbf{x}^i = \beta$ for all i and hence $\mathbf{x}^i \in F$ for all i.
- (2) $\Rightarrow$ (3): Obvious.
- (3) $\Rightarrow$ (4): If $\mathbf{c} \in \mathbb{R}^n$ such that $\max\{\mathbf{cx} : \mathbf{x} \in P\}$ is the finite number β, then the optimal face $F = \{\mathbf{x} \in P : \mathbf{cx} = \beta\}$ contains a minimal face which in turn contains an integral vector by (3).
- (4) $\Rightarrow$ (5): Obvious.
- (5) $\Rightarrow$ (6): Let $H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{ax} = \beta\}$ support P with $\mathbf{a} \in \mathbb{Q}^n$ and $\beta \in \mathbb{Q}$. We may scale $\mathbf{a}$ and β so that $\mathbf{a} \in \mathbb{Z}^n$ and $\mathbf{a}$ is primitive. Since H supports P, we may also assume that $\max\{\mathbf{ax} : \mathbf{x} \in P\} = \beta$. By (5), β is an integer and since g.c.d.$(a_j) = 1$ divides β, $\mathbf{ax} = \beta$ has an integer solution. (If g.c.d$(a_j) = 1$ then there exists integers $\{z_j\}$ such that $\sum a_j z_j = 1$. If β is an integer then $\sum a_j(\beta z_j) = \beta$.)
- (6) $\Rightarrow$ (3): We will prove the contrapositive. Suppose F is a minimal face of P without integral points. Let $A'\mathbf{x} \leq \mathbf{b}'$ be a subsystem of $A\mathbf{x} \leq \mathbf{b}$ such that $F = \{\mathbf{x} \in P : A'\mathbf{x} = \mathbf{b}'\}$. Since F is a minimal face, in fact, $F = \{\mathbf{x} \in \mathbb{R}^n : A'\mathbf{x} = \mathbf{b}'\}$. If $A'\mathbf{x} = \mathbf{b}'$ does not have an integer solution, then there exists a rational $\mathbf{y}$ such $\mathbf{y}A' \in \mathbb{Z}^n$ but $\mathbf{yb}' \notin \mathbb{Z}$. (This is an alternative theorem for the feasibility of $A\mathbf{x} = \mathbf{b}$, $\mathbf{x}$ integer that we will prove later.) Since A and $\mathbf{b}$ are integral, the property that $\mathbf{y}A' \in \mathbb{Z}^n$ but $\mathbf{yb}' \notin \mathbb{Z}$ remains so if we replace $\mathbf{y}$ by $\mathbf{y} + \mathbf{z}$ where $\mathbf{z} \in \mathbb{Z}^n$. Hence we may assume that $\mathbf{y} > \mathbf{0}$. Let $\mathbf{a} := \mathbf{y}A'$ and $\beta := \mathbf{yb}'$. Then $H = \{\mathbf{x} : \mathbf{ax} = \beta\}$ supports P. To see this, check that $\mathbf{x} \in P \Rightarrow A\mathbf{x} \leq \mathbf{b} \Rightarrow A'\mathbf{x} \leq \mathbf{b}' \Rightarrow \mathbf{y}A'\mathbf{x} \leq \mathbf{yb}' \Rightarrow \mathbf{ax} \leq \beta$ and $\mathbf{x} \in F \Rightarrow A'\mathbf{x} = \mathbf{b}' \Rightarrow \mathbf{ax} = \beta$. But H cannot contain any integer points since $\mathbf{a} \in \mathbb{Z}^n$ and $\beta \notin \mathbb{Z}$. Therefore, (6) fails.
- (3) $\Rightarrow$ (1): Let F_i, $1 \leq i \leq p$ be the minimal faces of P. For each F_i, choose an integer point $\mathbf{x}^i$ on it. Let $Q := \mathrm{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ and K be the recession cone of P. Then
$$P = Q + K = Q^I + K^I \subseteq (Q + K)^I = P^I \subseteq P.$$

This implies that $P = P^I$.

$\square$

See Remark 12.17 for a one line proof of (1) $\Leftrightarrow$ (6) that follows from a non-trivial theorem about integer hulls.

REMARK 7.6. (1) If P is pointed, then $P = P^I$ if and only if every vertex of P is integral.

(2) The convex hull of a finite set of integer points in $\mathbb{R}^n$ is an integer polytope (sometimes also called a **lattice polytope**). If the lattice points lie in $\{0,1\}^n$ then we call this lattice polytope a $0/1$**-polytope**.
(3) Every integer hull of a rational polyhedron is an integer polyhedron. The inequality description of this integer hull could be highly complicated.
(4) The basic integer program $\max\{\mathbf{cx} : \mathbf{x} \in P \cap \mathbb{Z}^n\}$ over a rational polytope P equals the linear program $\max\{\mathbf{cx} : \mathbf{x} \in P^I\}$ over the integer hull of P if we can find an inequality description of the integer polytope P^I. However, this latter task is difficult as we will see later creating the great divide between the complexity of an integer program and a linear program. If $P = P^I$ then every bounded integer program $\max\{\mathbf{cx} : \mathbf{x} \in P \cap \mathbb{Z}^n\}$ can be solved in polynomial time.

EXERCISE 7.7. (**HW**)

(1) Let $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$. Verify using the computer or otherwise that the polytope $P = \{\mathbf{x} \in \mathbb{R}^4 : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ is integral.
(2) More generally, let $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & p & q & r \end{pmatrix}$ where $0 < p < q < r$, $p, q, r \in \mathbb{N}$, $\gcd(p, q, r) = 1$ and $\mathbf{b} = \begin{pmatrix} r + q - p \\ qr \end{pmatrix}$. Prove that $P = \{\mathbf{x} \in \mathbb{R}^4 : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ is integral.

Most combinatorial optimization problems are integer programs over 0/1-vectors. These problems are typically formulated in the form $\max\{\mathbf{cx} : \mathbf{x} \in P \cap \mathbb{Z}^n\}$ where the convex hull of the feasible 0/1-vectors is the integer hull of P.

EXERCISE 7.8. (**HW**) Show that every 0/1-polytope in $\mathbb{R}^n$ can be expressed as the integer hull P^I of a rational polytope P in the unit cube $\{\mathbf{x} \in \mathbb{R}^n : 0 \leq x_i \leq 1 \;\; \forall \; i\}$ where P is given in the form $A\mathbf{x} \leq \mathbf{b}$ with every entry of A one of $0, 1$ or -1.

EXERCISE 7.9. (**HW**)

(1) Let $G = (V, E)$ be an undirected graph and $M(G)$ its matching polytope. Prove that $M(G)$ is the integer hull of the polytope $P(G) \subset \mathbb{R}^{|E|}$ described by the inequalities:
$$x_e \geq 0 \;\; \forall \; e \in E, \quad \sum_{v \in e} x_e \leq 1 \;\; \forall \; v \in V.$$
(2) Calculate the matching polytope of the complete graph K_4. Find all the facet inequalities needed in the matching polytope that were not present in the description of $P(G)$.

CHAPTER 8

Total Unimodularity and Unimodularity

Given a square submatrix B, of a matrix $A \in \mathbb{Z}^{m\times n}$, we call $\det(B)$ a **minor** of A. The maximal minors of A come from the maximal square submatrices in A.

Definition 8.1. An integral $m \times n$ matrix A of full row rank is **unimodular** if each non-zero maximal minor of A is ± 1.

Definition 8.2. A matrix A is **totally unimodular** if each minor of A is $0, 1$ or -1.

Exercise 8.3. Prove that A is totally unimodular if and only if the matrix $[I\ A]$ is unimodular.

Totally unimodular matrices give rise to integral polyhedra.

Theorem 8.4. *If $A \in \mathbb{Z}^{m\times n}$ is totally unimodular and $\mathbf{b} \in \mathbb{Z}^m$ then the polyhedron $P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ is integral.*

Proof. Let $F = \{\mathbf{x} : A'\mathbf{x} = \mathbf{b}'\}$ be a minimal face of P where $A'\mathbf{x} \le \mathbf{b}'$ is a subsystem of $A\mathbf{x} \le \mathbf{b}$. We may assume that A' has full row rank and of the form $A' = [U\, V]$ where U is a non-singular matrix. Since A is totally unimodular, $\det(U) = \pm 1$. Hence F contains

$$\mathbf{x} = \begin{pmatrix} U^{-1}\mathbf{b}' \\ \mathbf{0} \end{pmatrix} \in \mathbb{Z}^n.$$

□

Theorem 8.5. *Let A be an integral matrix of full row rank. Prove that the polyhedron $P = \{\mathbf{x} : \mathbf{x} \ge \mathbf{0},\ A\mathbf{x} = \mathbf{b}\}$ is integral for each integral vector $\mathbf{b}$ if and only if A is unimodular.*

Proof. Suppose A is unimodular and $\mathbf{b}$ is integral. Check that the lineality space of $P = \{0\}$ and hence every minimal face of P is a vertex. Let $\mathbf{x}'$ be a vertex of P. Then there exists n linearly independent contrainsts that hold at equality at $\mathbf{x}'$. Let A' be the submatrix of A whose columns are indexed by the non-zero entries in $\mathbf{x}'$. If the columns of A' are linearly dependent then there exists a $0 \ne \mathbf{y}$ such that $A'\mathbf{y} = \mathbf{0}$. Then $\mathbf{x}' \pm \epsilon\mathbf{y}$ satisfies $A\mathbf{x} = \mathbf{b}$ and all the inequalities from $\mathbf{x} \ge 0$ that held at equality at $\mathbf{x}'$. This would mean that the minimal face containing $\mathbf{x}'$ is not zero-dimensional which is false. Therefore, the columns of A' are linearly independent and we can extend them to a maximal non-singular submatrix B of A. Then $\mathbf{x}' = (B^{-1}\mathbf{b}, \mathbf{0})$ and since $\det(B) = \pm 1$, $\mathbf{x}'$ is integral.

Conversely, suppose P is integral whenever $\mathbf{b}$ is integral. Let B be a maximal non-singular submatrix of A. We need to argue that B has determinant ± 1. Note that $\det(B) = \pm 1$ if and only if $B^{-1}\mathbf{t}$ is integral for all $\mathbf{t}$ integral. (Why?) So let $\mathbf{t}$ be an integral vector. Pick $\mathbf{y}$ integral so that $\mathbf{z} := \mathbf{y} + B^{-1}\mathbf{t} \ge 0$. Then $\mathbf{b} = B\mathbf{z} = B\mathbf{y} + \mathbf{t}$ is integral. Extend $\mathbf{z}$ to $\mathbf{z}'$ by adding zero components so that $A\mathbf{z}' = B\mathbf{z} = \mathbf{b}$. Then $\mathbf{z}'$ lies in $\{\mathbf{x} : \mathbf{x} \ge 0, A\mathbf{x} = \mathbf{b}\}$. If $A \in \mathbb{Z}^{m\times n}$ of rank m, then $\mathbf{z}'$ satisfies n linearly independent constraints from among

$A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0$ at equality (m constraints from $A\mathbf{x} = \mathbf{b}$ and at least $n - m$ of the $x_i \geq 0$ constraints.) Therefore, $\mathbf{z}'$ is a vertex of P and hence integral by assumption. This implies that $\mathbf{z}$ is also integral and hence $B^{-1}\mathbf{t} = \mathbf{z} - \mathbf{y}$ is integral. □

Corollary 8.6. *A matrix A is totally unimodular if and only if for each integral vector $\mathbf{b}$ the polyhedron $\{\mathbf{x} : \mathbf{x} \geq \mathbf{0},\ A\mathbf{x} \leq \mathbf{b}\}$ is integral.*

Proof. Recall that A is totally unimodular if and only if $[I\ A]$ is unimodular. By Theorem 8.5 $[I\ A]$ is unimodular if and only if for every integral $\mathbf{b}$,

$$\{(\mathbf{y}, \mathbf{x}) : (\mathbf{y}, \mathbf{x}) \geq \mathbf{0}, I\mathbf{y} + A\mathbf{x} = \mathbf{b}\} = \{\mathbf{x} : \mathbf{x} \geq \mathbf{0}, A\mathbf{x} \leq \mathbf{b}\}$$

is integral. □

One can check that a matrix is totally unimodular in many different ways. Below we summarize some of these characterizations. See [**Sch86**] for proofs and for more characterizations.

Theorem 8.7 (Theorem 19.3 [**Sch86**]). *Let $A \in \{0, \pm 1\}^{m \times n}$. Then the following are equivalent.*

(1) *A is totally unimodular.*
(2) *For each integral $\mathbf{b}$, the polyhedron $\{\mathbf{x} : \mathbf{x} \geq \mathbf{0},\ A\mathbf{x} \leq \mathbf{b}\}$ is integral.*
(3) *For all integer vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$, the polyhedron $\{\mathbf{x} : \mathbf{d} \leq \mathbf{x} \leq \mathbf{c},\ \mathbf{a} \leq A\mathbf{x} \leq \mathbf{b}\}$ is integral.*
(4) *Each collection of columns of A can be split into two parts so that the sum of the columns in one part minus the sum of the columns in the other part is a vector with entries only $0, \pm 1$.*
(5) *Each non-singular submatrix of A has a row with an odd number of non-zero components.*
(6) *No square submatrix of A has determinant ± 2.*

Exercise 8.8. (**HW**) Prove that the following conditions are equivalent for a matrix A of full row rank.

(1) For each maximal non-singular submatrix B of A, the matrix $B^{-1}A$ is integral.
(2) For each maximal non-singular submatrix B of A, the matrix $B^{-1}A$ is totally unimodular.
(3) There exists a maximal non-singular submatrix B of A for which $B^{-1}A$ is totally unimodular.

Exercise 8.9. (**HW**) Prove that A is totally unimodular if and only if for each non-singular submatrix B of A and each non-zero $\{0, \pm 1\}$-vector $\mathbf{y}$, the gcd of the entries in $\mathbf{y}B$ is one.

Example 8.10.
(1) *Unidirected bipartite graphs*: Let $G = (V, E)$ be an undirected graph and $M(G)$ be its vertex-edge incidence matrix. This means that for $v \in V$ and $e \in E$, the (v, e)-entry of $M(G)$ is one if $v \in e$ and zero otherwise. Then $M(G)$ is totally unimodular if and only if G is bipartite.
(2) *Directed graphs*: Let $D = (V, A)$ be a directed graph and $M(D)$ be the $V \times A$ incidence matrix of D described as follows. For $v \in V$ and $a \in A$, the (v, a)-entry of $M(D)$ is 1 if v is the head of a, -1 if v is the tail of a and 0 otherwise. The matrix $M(D)$ is totally unimodular.

(3) *Network matrices*: Let $D = (V, A)$ be a directed graph and let $T = (V, A_0)$ be a directed spanning tree in D. Let M be the $A_0 \times A$ matrix defined as follows: for $a' = (v, w) \in A_0$ and $a \in A$,

$$M_{a',a} := \begin{cases} 1 \text{ if if the unique } v, w \text{ path in } T \text{ passes through } a' \text{ forwardly} \\ -1 \text{ if if the unique } v, w \text{ path in } T \text{ passes through } a' \text{ backwardly} \\ 0 \text{ if the unique } v, w \text{ path in } T \text{ does not pass through } a'. \end{cases}$$

In 1980, Seymour proved that all totally unimodular matrices can be built from network matrices and two other matrices by using eight types of operations. The decomposition of totally unimodular matrices given by his theorem can be described in polynomial time. This means that the problem of testing deciding whether a matrix is totally unimodular lies in $NP \cap co - NP$. Seymour also gave a polynomial time algorithm to test whether a matrix is totally unimodular but this is quite complicated. The upshot of all this is that given an integer matrix A one can decide in polynomial time whether the polyhedron $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ or $P = \{\mathbf{x} : \mathbf{x} \geq \mathbf{0}, A\mathbf{x} \leq \mathbf{b}\}$ is integral for every $\mathbf{b}$ without having to compute P^I.

CHAPTER 9

Applications of Total Unimodularity

In this chapter we will see a few applications of total unimodularity.

DEFINITION 9.1. Let $G = (V, E)$ be an undirected graph.

(1) A **stable set** in G is a set $S \subseteq V$ such that for every pair $v, w \in S$, $\{v, w\} \notin E$.
(2) A **vertex cover** in G is a set $W \subseteq V$ such that for every $e \in E$, there exists some $v \in W$ such that $v \in e$.
(3) A **edge cover** in G is a set $F \subseteq E$ such that for every $v \in V$, $v \in f$ for some $f \in F$.

Note that S is a stable set in G with characteristic vector $\mathbf{y} := \chi^S$ if and only if $\mathbf{y} \in \{0, 1\}^V$ and $\mathbf{y}M \leq \mathbf{1}$. This in turn is equivalent to $\mathbf{y} \geq \mathbf{0}, \mathbf{y}M \leq \mathbf{1}$, $\mathbf{y}$ integral.

Similarly, F is an edge cover of G with characteristic vector $\mathbf{x} := \chi^F$ if and only if $\mathbf{x} \in \{0, 1\}^E$ and $M\mathbf{x} \geq \mathbf{1}$.

Also note that the maximum size of a stable set in G is less than or equal to the minimum size of an edge cover in G. If G is bipartite, then you get equality of this max and min.

THEOREM 9.2 (König 1933). *The maximum size of a stable set in a bipartite graph G equals the minimum size of an edge cover in the graph.*

PROOF. Let G be a bipartite graph. Then its incidence matrix M is totally unimodular. By the observations above, the maximum size of a stable set in G is the value of the integer program

$$\max\{\mathbf{1} \cdot \mathbf{y} \,:\, \mathbf{y} \geq \mathbf{0},\, \mathbf{y}M \leq \mathbf{1},\, \mathbf{y}\,\text{integral}\}$$

Since M is totally unimodular, this integer program equals the linear program

$$\max\{\mathbf{1} \cdot \mathbf{y} \,:\, \mathbf{y} \geq \mathbf{0},\, \mathbf{y}M \leq \mathbf{1}\}$$

which by LP-duality equals the linear program

$$\min\{\mathbf{1} \cdot \mathbf{x} \,:\, \mathbf{x} \geq \mathbf{0},\, M\mathbf{x} \geq \mathbf{1}\}.$$

Again using the total unimodularity of M this min LP equals the integer program

$$\min\{\mathbf{1} \cdot \mathbf{x} \,:\, \mathbf{x} \geq \mathbf{0},\, M\mathbf{x} \geq \mathbf{1},\, \mathbf{x}\,\text{integer}\}.$$

Note that the optimal value of this last integer program is precisely the minimum size of an edge cover in G. $\square$

Similarly, one gets the following relation which leads to the theorem following it.

$$\max\{\mathbf{1} \cdot \mathbf{x} \,:\, \mathbf{x} \geq \mathbf{0},\, M\mathbf{x} \leq \mathbf{1},\, \mathbf{x}\,\text{integer}\} = \min\{\mathbf{1} \cdot \mathbf{y} \,:\, \mathbf{y} \geq \mathbf{0},\, \mathbf{y}M \geq \mathbf{1},\, \mathbf{y}\,\text{integral}\}.$$

THEOREM 9.3 (König, Egerváry 1931). *The maximum size of a matching in a bipartite graph equals the minimum size of a vertex cover in the graph.*

EXERCISE 9.4. (**HW**)

(1) Write down inequality descriptions of the matching polytope and the perfect matching polytope of a bipartite graph G (using M).
(2) Prove that the perfect matching polytope of the complete bipartite graph $K_{n,n}$ equals the Birkhoff polytope $B(n)$.

We now turn out attention to directed graphs. Let $D = (V, A)$ be a directed graph. Then we saw earlier that its node-arc incidence matrix M is totally unimodular. Let $\mathbf{x} := (x_a : a \in A) \in \mathbb{R}^n_{\geq 0}$ such that $M\mathbf{x} = \mathbf{0}$. Then $\mathbf{x}$ is called a **circulation** in G since for every $v \in V$,

$$\sum_{a \text{ enters } v} x_a = \sum_{a \text{ leaves } v} x_a.$$

In other words, $\mathbf{x}$ is a **flow** in G that is conserved at every vertex. By Theorem 8.7 (3), for integral vectors $\mathbf{c}$ and $\mathbf{d}$,

$$\{\mathbf{x} : \mathbf{d} \leq \mathbf{x} \leq \mathbf{c}, \mathbf{0} \leq M\mathbf{x} \leq \mathbf{0}\} = \{\mathbf{x} : \mathbf{d} \leq \mathbf{x} \leq \mathbf{c}, M\mathbf{x} = \mathbf{0}\}$$

is integral. In other words, if there is a circulation $\mathbf{x}$ in D with $\mathbf{d} \leq \mathbf{x} \leq \mathbf{c}$ then there is an integral such circulation.

Now pick a cost vector $\mathbf{f} = (f_a : a \in A)$ such that each $f_a \in \mathbb{Z} \cup \{\pm\infty\}$ and consider the integer program

$$\max\{\mathbf{f} \cdot \mathbf{x} : \mathbf{d} \leq \mathbf{x} \leq \mathbf{c}, M\mathbf{x} = \mathbf{0}, \mathbf{x} \,\text{integral}\}.$$

Since M is totally unimodular, this integer program is equivalent to the linear program

$$\max\{\mathbf{f} \cdot \mathbf{x} : \mathbf{d} \leq \mathbf{x} \leq \mathbf{c}, M\mathbf{x} = \mathbf{0}\}$$

which by LP-duality and the total unimodularity of M is equal to the integer program

$$\min\{\mathbf{y}'\mathbf{c} - \mathbf{y}''\mathbf{d} : \mathbf{z}M + \mathbf{y}' - \mathbf{y}'' = \mathbf{f},\ \mathbf{y}', \mathbf{y}'' \in \mathbb{N}^A, \mathbf{z} \in \mathbb{Z}^V\}.$$

We now specialize this integer program in several different ways to derive well-known results in optimization and combinatorics.

(1) Pick an arc $a_0 := (s, r)$ in D and let $f_{a_0} = 1$. For all $a \in A$, $a \neq a_0$, let $f_a = 0$. Further, let $\mathbf{d} = \mathbf{0}$ and choose $\mathbf{c} \geq \mathbf{0}$ with $c_{a_0} = \infty$. Then the max integer program above becomes

$$\max\{x_{a_0} : 0 \leq x_a \leq c_a \ \forall\, a \neq a_0, 0 \leq x_{a_0}, M\mathbf{x} = \mathbf{0}, \mathbf{x} \,\text{integral}\}$$

which is maximizing the flow from s to r in D subject to the capacity constraints on the arcs given by $\mathbf{c}$. But this is equivalent (by flow conservation at each node) to maximizing the $r - s$-flow in the digraph $D' = (V, A \backslash \{a_0\})$ subject to the capacity constraint c_a on each arc $a \in A \backslash \{a_0\}$. In other words, this max integer program is the same as the linear program

$$\begin{array}{lll} \max & \sum_{a \text{ leaves } r} x_a - \sum_{a \text{ enters } r} x_a & \\ \text{s.t.} & \sum_{a \text{ leaves } v} x_a = \sum_{a \text{ enters } v} x_a & \forall\, v \in V \backslash \{r, s\} \\ & 0 \leq x_a \leq c_a & \forall\, a \in A \backslash \{a_0\} \end{array}.$$

Specializing the equivalent min integer program as well we have

$$\min\{\mathbf{y}'\mathbf{c} : \exists\, \mathbf{z} \in \mathbb{Z}^V \text{ s.t. } y'_a - z_v + z_u \geq 0 \text{ if } a = (u, v) \neq (s, r), y'_{a_0} - z_r + z_s \geq 1, \mathbf{y}' \in \mathbb{N}^A\}.$$

Since $c_{a_0} = \infty$, in the optimal solution of this min integer program we will have $y'_{a_0} = 0$. Therefore, the last constraint can be replaced by $z_s \geq z_r + 1$.

Let $U := \{v \in V \,:\, z_v \geq z_s\}$ and let $\delta^{-1}(U)$ be the set of arcs in A entering U. Note that $s \in U$ but $r \notin U$ since $z_s > z_r$. Therefore U and $V \backslash U$ create an $r-s$-cut. If $a = (u,v) \in \delta^{-1}(U)$ then $z_v \geq z_s$ by definition of U and since $u \notin U$, $z_u < z_s$. Therefore, $y'_a = z_v - z_u \geq 1$ since both z_u and z_v are integers. Thus the optimal cost value $\mathbf{y}'\mathbf{c} \geq \sum(c_a \,:\, a \in \delta^{-1}(U))$. In other words, the maximum $r-s$-flow in D' is at least as large as the capacity of the $r-s$-cut $\delta^{-1}(U)$. But the capacity of any $r-s$-cut is also an upper bound to the maximum $r-s$-flow in D' which leads us to the following famous theorem in optimization.

Theorem 9.5 (Ford-Fulkerson 1956). *The maximum value of an $r-s$-flow subject to capacity $\mathbf{c}$ equals the minimum capacity of any $r-s$-cut. Further, if all the capacities are integers, then the optimal flow is integral.*

(2) Take $c_a = 1$ for all $a \neq a_0$ and consider the optimal (max) $r-s$-flow, say $\mathbf{x}$. Then $\mathbf{x}$ is integeral and on each arc a that carries flow, $x_a = 1$. For this flow, a directed path from r to s with non-zero flow cannot intersect another such path at an arc due to flow conservation and capacity limits. Thus the maximum $r-s$-flow is also the maximum number of arc-disjoint $r-s$-paths in the graph.

Theorem 9.6 (Menger 1927). *Given $D = (V,A)$ and vertices r,s, the maximum number of pairwise arc disjoint $r-s$-paths in D equals the minimum size of of any $r-s$-cut.*

(3) Here is a third theorem that also follows from specializing the above pair of integer programs.

Theorem 9.7 (Dilworth 1950). *In a partially ordered set $(X, \leq)$, the maximum size of an antichain (set of pairwise incomparable elements) equals the minimum number of chains needed to cover X.*

A chain in X is a subset of the form $x_1 \leq x_2 \leq \ldots$.

Exercise 9.8. (**HW**) Explain how Dilworth's theorem is a third specialization of the min-max integer programs used to prove the above two theorems.

CHAPTER 10

Hilbert Bases

DEFINITION 10.1. A finite set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_t$ is a **Hilbert basis** if every integral vector $\mathbf{b}$ in $\text{cone}(\mathbf{a}_1, \ldots, \mathbf{a}_t)$ is a non-negative integral combination of $\mathbf{a}_1, \ldots, \mathbf{a}_t$.

If $\mathbf{a}_1, \ldots, \mathbf{a}_t$ is a Hilbert basis, we often refer to it as a Hilbert basis of $\text{cone}(\mathbf{a}_1, \ldots, \mathbf{a}_t)$.

EXAMPLE 10.2. The vectors $(1,0), (1,1), (1,2), \ldots, (1,k) \in \mathbb{N}^2$ form a Hilbert basis. However, any subset of the above set that leaves out one of the vectors $(1,j)$, $1 \leq j \leq k-1$ is not a Hilbert basis. Note that $\text{cone}((1,0),(1,1),(1,2),\ldots,(1,k))$ is spanned by $(1,0)$ and $(1,k)$. Therefore, the input to this Hilbert basis calculation has size $\mathcal{O}(\log k)$ while the output has size $\mathcal{O}(k \log k)$. Therefore, Hilbert bases computations cannot be done in polynomial time in the input size as the output size can be exponentially larger than the input size.

THEOREM 10.3. [**Sch86**, Theorem 16.4] *Every rational polyhedral cone C is generated by an integral Hilbert basis. If C is pointed there is a unique minimal integral Hilbert basis generating C.*

PROOF. Let $\mathbf{c}_1, \ldots, \mathbf{c}_k$ be primitive integral vectors that generate C. Consider the parallelepiped

$$Z = \{\sum_{i=1}^{k} \mu_i \mathbf{c}_i \,:\, 0 \leq \mu_i \leq 1\}.$$

We prove that the set H of integral vectors in Z form a Hilbert basis of C. Since $\mathbf{c}_1, \ldots \mathbf{c}_k \in Z$, H generates C. Suppose $\mathbf{c}$ is any integral vector in C. Then $\mathbf{c} = \sum_{j=1}^{k} \lambda_j \mathbf{c}_j$ where $\lambda_j \geq 0$. Rewrite as

$$\mathbf{c} = \sum_{j=1}^{k} (\lfloor \lambda_j \rfloor + (\lambda_j - \lfloor \lambda_j \rfloor)) \mathbf{c}_j$$

and then again as

$$\mathbf{c} - \sum_{j=1}^{k} \lfloor \lambda_j \rfloor \mathbf{c}_j = \sum_{j=1}^{k} (\lambda_j - \lfloor \lambda_j \rfloor) \mathbf{c}_j.$$

Since the left-hand-side is integral, so is the right-hand-side. However, the right-hand side belongs to Z and hence $\mathbf{c}$ is a non-negative integer combination of elements in H. This proves that H is a Hilbert basis.

Now suppose C is pointed. Then there exists a vector $\mathbf{b}$ such that $\mathbf{bx} > 0$ for all $\mathbf{x} \in C \backslash \{\mathbf{0}\}$. Let

$$H' = \{\mathbf{a} \in C \,:\, \mathbf{a} \neq \mathbf{0},\, \mathbf{a} \text{ integral}, \ \mathbf{a} \text{ not the sum of two other integral vectors in } C\}.$$

The set H' is finite since it must be contained in any integral Hilbert basis of C. If H' is not a Hilbert basis of C, then we can choose an integral $\mathbf{c} \in C$ such that $\mathbf{c} \notin \mathbb{N}H'$ and $\mathbf{bc}$ is as small as possible. Note that this $\mathbf{c}$ must be in Z and hence there is one that minimizes $\mathbf{bx}$ over Z. Then since $\mathbf{c} \notin H'$, there exists $\mathbf{c}_1$ and $\mathbf{c}_2$ non-zero integral vectors in C such that $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2$. Since $\mathbf{bc}, \mathbf{bc}_1, \mathbf{bc}_2$ are all positive and $\mathbf{bc} = \mathbf{bc}_1 + \mathbf{bc}_2$, we get that both $\mathbf{bc}_1$ and $\mathbf{bc}_2$ are less than $\mathbf{bc}$. By assumption then $\mathbf{c}_1$ and $\mathbf{c}_2$ lie in $\mathbb{N}H'$ and hence so does $\mathbf{c} = \mathbf{c}_1 + \mathbf{c}_2$, a contradiction. □

REMARK 10.4.
(1) Note that if C is not pointed then there is no unique minimal integral Hilbert basis. For instance if $C = \mathbb{R}^2$, then $\pm(1,0), \pm(0,1)$ form a minimal Hilbert basis. But so does $(1,0), (0,1), (-1,-1)$.
(2) Note that every element in a minimal Hilbert basis is primitive.

EXAMPLE 10.5. The software package **Normaliz** [**BK**] can be used to compute Hilbert bases of rational cones. For instance, suppose we wish to compute the unique minimal Hilbert basis of the pointed cone in $\mathbb{R}^4$ generated by $(1,0,0,0), (0,1,0,0), (0,0,1,0)$ and $(1,2,3,4)$. Then **Normaliz** is used as follows.

```
[thomas@rosa]more example.in
4          --- number of generators of cone
4          --- dimension of the vectors
1 0 0 0    --- the four vectors row-wise
0 1 0 0
0 0 1 0
1 2 3 4
0          --- computes Hilbert basis wrt the ambient integer lattice

[thomas@rosa] normaliz example
[thomas@rosa] more example.out
7 generators of integral closure:  --- Hilbert basis has 7 elements
  1  0  0  0
  0  1  0  0
  0  0  1  0
  1  2  3  4
  1  2  3  3
  1  1  2  2
  1  1  1  1

(original) semigroup has rank 4 (maximal)
(original) semigroup is of index 4

4 support hyperplanes:
   0   0   0   1
   0   0   4  -3
   0   2   0  -1
   4   0   0  -1

(original) semigroup is not homogeneous
```

DEFINITION 10.6. A rational polyhedral pointed cone in $\mathbb{R}^n$ is **unimodular** if it has at most n extreme rays and the set of primitive integral vectors generating the extreme rays form part of a basis for $\mathbb{Z}^n$.

Hilbert bases play a crucial role in the theory of lattice points in polyhedra and in integer programming. They are so important that they have been studied in their own right. Two such questions were whether the cone generated by a Hilbert basis admits triangulations or covers by unimodular subcones that are generated by the elements of the Hilbert basis. All Hilbert bases in $\mathbb{R}^3$ admit unimodular triangulations but not in $\mathbb{R}^4$ and above. For instance, the seven element Hilbert basis in Example 10.5 does not admit a unimodular triangulation [**FZ99**]. Similarly, unimodular covers do not exist in $\mathbb{R}^6$ and above [**BG99**]. The cover question is open in $\mathbb{R}^4$ and $\mathbb{R}^5$.

EXAMPLE 10.7. The cone generated by $(1, 2)$ and $(0, 1)$ is unimodular while the cone generated by $(1, 2)$ and $(1, 0)$ is not.

EXERCISE 10.8. (**HW**)

(1) Let $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^2$ be two linearly independent vectors and let C be the cone they span in $\mathbb{R}^2$. Let $\mathbf{h}_1 := \mathbf{u}, \mathbf{h}_2, \ldots, \mathbf{h}_{t-1}, \mathbf{h}_t := \mathbf{v}$ be the elements in the unique minimal Hilbert basis of C in cyclic order from $\mathbf{u}$ to $\mathbf{v}$. Prove that the cones $\mathrm{cone}(\mathbf{h}_i, \mathbf{h}_{i+1})$ are unimodular for $i = 1, \ldots, t-1$.
(2) Let C be a pointed rational polyhedral cone in $\mathbb{R}^2$ generated by $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^2$ and let C' be the convex hull of all the non-zero lattice points in C. Prove that the elements of the minimal Hilbert basis of C are precisely the lattice points that lie on the bounded part of the boundary of C' between $\mathbf{u}$ and $\mathbf{v}$. Give an example in $\mathbb{R}^3$ to show that this result does not hold in $\mathbb{R}^3$.

LEMMA 10.9. *Let $C = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{0}\}$ be a cone where A is an integral matrix with all subdeterminants of absolute value at most Δ. If $\mathbf{h}$ is in the fundamental parallelepiped spanned by a set of primitive generators of C, then $||\mathbf{h}||_\infty \leq n\Delta$.*

PROOF. We first show that it is possible to find integral generators of C whose infinity-norms are at most Δ. Recall that every such generator is a solution to some subsystem $A'\mathbf{x} = \mathbf{0}$ of the system $A\mathbf{x} \leq \mathbf{0}$. Assume that A' has full row rank. This rank is less than or equal to $n-1$. Also assume that $A' = [U\ V]$ where U is non-singular. Split $\mathbf{x}$ as $\mathbf{x} = (\mathbf{x}_U, \mathbf{x}_V)$. Then to solve for $A'\mathbf{x} = U\mathbf{x}_U + V\mathbf{x}_V = \mathbf{0}$, we can set $\mathbf{x}_V$ to any arbitrary value. Assume we set $\mathbf{x}_V$ to a $0,1$-vector with a 1 in the k-th position. Then we have $U\mathbf{x}_U = -\mathbf{v}_k$ where $\mathbf{v}_k$ is the k-th column of V. Applying Cramer's rule, we can find a solution $\mathbf{x}_U$ of this square system with i-th component a ratio of two subdeterminants of A' and hence A. All denominators of the components of $\mathbf{x}_U$ are $\det(U)$ and all numerators are at most Δ in absolute value. Clearing the denominator, we get an integer $\mathbf{x} = (\mathbf{x}_{U'}, \mathbf{x}_{V'})$ where all components are at most Δ in absolute value.

Suppose $\mathbf{h}$ is in the fundamental parallelepiped spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_s$. Then using

THEOREM 10.10 (Caratheodory). *Every element of a d-dimensional rational polyhedral cone lies in a subcone generated by d linearly independent generators of the cone.*

there exists n linearly independent vectors among the $\mathbf{u}_i$'s such that $\mathbf{h}$ is also in the fundamental parallelepiped spanned by these n vectors, say $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Therefore, there exists $0 \leq \lambda_1, \ldots, \lambda_n \leq 1$ such that $\mathbf{h} = \lambda_1\mathbf{u}_1 + \cdots + \lambda_n\mathbf{u}_n$. This implies that

$$|\mathbf{h}_j| \leq \sum_{i=1}^{n} \lambda_i |\mathbf{u}_{ij}| \leq (\sum_{i=1}^{n} \lambda_i)\Delta \leq n\Delta.$$

$\square$

We now give an application of Hilbert bases to solving integer programs. The statement of the theorem comes from Theorem 17.3 in [**Sch86**]. Our proof here is slightly different.

Theorem 10.11. [**Sch86**, Theorem 17.3] *Let $A \in \mathbb{Z}^{m\times n}$ with all subdeterminants at most Δ in absolute value and let $\mathbf{b} \in \mathbb{Z}^m$ and $\mathbf{c} \in \mathbb{R}^n$. Let $\mathbf{z}$ be a feasible but not optimal solution of the integer program $\max\{\mathbf{cx} \,:\, A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\}$. Then there exists a feasible solution $\mathbf{z}'$ such that $\mathbf{cz}' > \mathbf{cz}$ and $||\mathbf{z} - \mathbf{z}'||_\infty \le n\Delta$.*

Proof. Since $\mathbf{z}$ is not optimal, there exists some $\mathbf{z}'' \in \mathbb{Z}^n$ such that $A\mathbf{z}'' \le \mathbf{b}$ and $\mathbf{cz}'' > \mathbf{cz}$. Split $A\mathbf{x} \le \mathbf{b}$ into two subsystems $A_1\mathbf{x} \le \mathbf{b}_1$ and $A_2\mathbf{x} \le \mathbf{b}_2$ such that $A_1\mathbf{z} \le A_1\mathbf{z}''$ and $A_2\mathbf{z} \ge A_2\mathbf{z}''$. Let $C := \{\mathbf{u} \,:\, A_1\mathbf{u} \ge \mathbf{0},\ \ A_2\mathbf{u} \le \mathbf{0}\}$. Then $\mathbf{z}'' - \mathbf{z}$ is an integral vector in the cone C. Let $\mathbf{h}_1, \ldots, \mathbf{h}_t$ be a minimal Hilbert basis for C. Then

$$\mathbf{z}'' - \mathbf{z} = \sum_{i=1}^{t} \lambda_i \mathbf{h}_i$$

for some $\lambda_1, \ldots, \lambda_t \in \mathbb{N}$. Since $0 < \mathbf{c}(\mathbf{z}'' - \mathbf{z}) = \sum_{i=1}^t \lambda_i \mathbf{ch}_i$, there exists an $\mathbf{h}_i$ with $\lambda_i \ge 1$ such that $\mathbf{ch}_i > 0$. Consider the vector $\mathbf{z}' := \mathbf{z} + \mathbf{h}_i$. Since $\mathbf{z}'' - \mathbf{z} = \sum_{i=1}^t \lambda_i \mathbf{h}_i$, we have

$$\mathbf{z}'' - \sum_{j\neq i} \lambda_j \mathbf{h}_j - (\lambda_i - 1)\mathbf{h}_i = \mathbf{z} + \mathbf{h}_i = \mathbf{z}'.$$

Then $A\mathbf{z}' = A(\mathbf{z}'' - \sum_{j\neq i} \lambda_j \mathbf{h}_j - (\lambda_i - 1)\mathbf{h}_i)$. Now using the fact that $A_1\mathbf{h}_k \ge \mathbf{0}$ for $k = 1, \ldots, t$, check that $A_1\mathbf{z}' = A_1(\mathbf{z}'' - \sum_{j\neq i} \lambda_j \mathbf{h}_j - (\lambda_i - 1)\mathbf{h}_i) \le \mathbf{b}_1$. Similarly, $A_2\mathbf{z}' = A_2(\mathbf{z} + \mathbf{h}_i) = A_2\mathbf{z} + A_2\mathbf{h}_i \le A_2\mathbf{z} \le \mathbf{b}_2$. Further, $\mathbf{cz}' = \mathbf{c}(\mathbf{z} + \mathbf{h}_i) > \mathbf{cz}$ since $\mathbf{ch}_i > 0$. Therefore, $\mathbf{z}'$ is a feasible solution to the integer program $\max\{\mathbf{cx} \,:\, A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\}$ that improves the cost value.

To finish the proof we have to argue that $||\mathbf{z}' - \mathbf{z}||_\infty = ||\mathbf{h}_i||_\infty \le n\Delta$. This follows from Lemma 10.9. $\square$

Theorem 10.11 says that every non-optimal solution to the integer program $\max\{\mathbf{cx} \,:\, A\mathbf{x} \le \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\}$ can be improved by another feasible solution to the program that is not too far from the first solution. Improving vectors for integer programs are known as **test sets**. The above theorem is a variant of a construction due to Jack Graver who first showed the existence of test sets for integer programming. We only look at test sets briefly here since our interest is not so much in studying them but in just using the above result to study integer hulls as in the following theorem. Theorem 10.12 strengthens Theorem 7.2.

Theorem 10.12. [**Sch86**, Theorem 17.4] *For each rational matrix A there exists an integral matrix M such that for each vector $\mathbf{b}$ there is a vector $\mathbf{d}$ such that*

$$\{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} \le \mathbf{b}\}^I = \{\mathbf{x} \in \mathbb{R}^n \,:\, M\mathbf{x} \le \mathbf{d}\}.$$

If A is integral and all subdeterminants of A have absolute value at most Δ then we can take all entries in M to be at most $n^{2n}\Delta^n$ in absolute value.

Proof. Assume A is integral with all subdeterminants of absolute value at most Δ. Let

$$L := \{\mathbf{u} \,:\, \exists\ \mathbf{y} \ge \mathbf{0} \,:\, \mathbf{y}A = \mathbf{u},\ \mathbf{u} \text{ integral},\ ||\mathbf{u}||_\infty \le n^{2n}\Delta^n\}.$$

In other words, L is the set of all integral vectors in the cone spanned by the rows of A with infinity norm at most $n^{2n}\Delta^n$. Let M be the matrix whose rows are all the elements of L. We will prove that M can be taken to be the matrix needed in the theorem. Recall that a vector $\mathbf{c}$ is bounded over the polyhedron $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ if and only if $\mathbf{c}$ lies in the polar of the recession cone $C = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$ of the polyhedron. The elements in L come from the polar of this recession cone and hence the linear functional $\mathbf{mx}$, for each row $\mathbf{m}$ of M, attains a finite maximum over every $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ as $\mathbf{b}$ varies and hence also over their integer hulls.

Fix $\mathbf{b}$. If $A\mathbf{x} \leq \mathbf{b}$ has no solution, then we can choose a $\mathbf{d}$ such that $M\mathbf{x} \leq \mathbf{d}$ also has no solution. (Note that the rows of A are among the rows of M and so we can choose $\mathbf{d}$ so that the right-hand-sides of the inequalities from the rows of A are b_i.)

If $A\mathbf{x} \leq \mathbf{b}$ is feasible but does not have an integral solution, then its recession cone $\{\mathbf{y} : A\mathbf{y} \leq \mathbf{0}\}$ is not full-dimensional which means that it has an implicit equality, say $\mathbf{ax} \leq 0$. Then both $\mathbf{a}$ and $-\mathbf{a}$ belongs to L and we can choose $\mathbf{d}$ so that $M\mathbf{x} \leq \mathbf{d}$ is infeasible.

So assume that $A\mathbf{x} \leq \mathbf{b}$ has an integral solution. For each vector $\mathbf{c} \in \mathbb{R}^n$ let

$$\delta_{\mathbf{c}} := \max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\}.$$

It suffices to show that

$$\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}^I = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{ux} \leq \delta_{\mathbf{u}}\ \forall\ \mathbf{u} \in L\}.$$

(By our earlier discussion, $\delta_{\mathbf{u}}$ is finite for all $\mathbf{u} \in L$.) Since we can think of the left-hand-side as the intersection of all half-spaces $\mathbf{cx} \leq \delta_{\mathbf{c}}$ as $\mathbf{c}$ varies over all the integer vectors in the polar of the recession cone $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{0}\}$, we get that the left-hand-side is contained in the right-hand-side. To show the opposite containment, let $\mathbf{cx} \leq \delta_{\mathbf{c}}$ be a valid inequality for $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}^I$. We will show that $\mathbf{cx} \leq \delta_{\mathbf{c}}$ is also valid for $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{ux} \leq \delta_{\mathbf{u}}\ \forall\ \mathbf{u} \in L\}$ which will prove the containment. Let $\mathbf{z}$ be an optimal solution of the integer program $\max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\}$. Consider the cones

$$K := \text{cone}\{\mathbf{x} - \mathbf{z} : A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n\} \text{ and}$$
$$K' := \text{cone}\{\mathbf{x} - \mathbf{z} : A\mathbf{x} \leq \mathbf{b},\ \mathbf{x} \in \mathbb{Z}^n,\ \ ||\mathbf{x} - \mathbf{z}||_\infty \leq n\Delta\}.$$

Clearly, $K' \subseteq K$.

Exercise 10.13. (**HW**)
(1) Using Theorem 10.11, prove that $K' = K$.
(2) Prove that $K = \{\mathbf{y} : \mathbf{uy} \leq 0,\ \mathbf{u} \in L_1\}$ for some subset L_1 of L.

For each $\mathbf{u} \in L_1$, $\delta_{\mathbf{u}} = \mathbf{uz}$. Also, for each $\mathbf{y} \in K$, $\mathbf{cy} \leq 0$ which implies that $\mathbf{c} \in K^*$. Since K^* is generated by all the vectors $\mathbf{u} \in L_1$, we get that

$$\mathbf{c} = \lambda_1 \mathbf{u}_1 + \cdots + \lambda_t \mathbf{u}_t$$

for some $\lambda_1, \ldots, \lambda_t \geq 0$ and $\mathbf{u}_1, \ldots, \mathbf{u}_t \in L_1$. Further, $\delta_{\mathbf{c}} = \mathbf{cz} = \lambda_1 \mathbf{u}_1 \mathbf{z} + \cdots + \lambda_t \mathbf{u}_t \mathbf{z} = \lambda_1 \delta_{\mathbf{u}_1} + \cdots + \lambda_t \delta_{\mathbf{u}_t}$ and hence the inequality $\mathbf{cx} \leq \delta_{\mathbf{c}}$ is valid for

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{ux} \leq \delta_{\mathbf{u}}\ \forall\ \mathbf{u} \in L\}.$$

$\square$

CHAPTER 11

An Integral Alternative Theorem

The goal of this lecture is to present an alternative theorem for the feasibility of a system of rational linear equations $A\mathbf{x} = \mathbf{b}$ over the integers. This leads to the result that the problem of deciding whether $A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}^n$ is feasible lies in both $\mathcal{NP}$ and $co - \mathcal{NP}$. This problem actually lies in $\mathcal{P}$ but we will not prove that here and will only allude to how such a result can be proved.

Definition 11.1. A matrix A of full row rank is in **Hermite normal form** if it has the form $[B\ 0]$ where B is a non-singular, lower triangular, non-negative matrix with a unique maximum entry in each row located on the main diagonal of B.

The following operations on the columns (or rows) of a matrix are called **elementary unimodular operations**.

(1) exchanging two columns,
(2) multiplying a column by a -1,
(3) adding an integral multiple of one column to another column.

The Hermite normal form of a matrix is the integer analog of the row-echelon form of a matrix from linear algebra obtained at the end of Gaussian elimination. In Gaussian elimination, we use operations that keep the row-space of the matrix (as a vector space) intact. In integer linear algebra, we want to keep the lattice spanned by the columns (or rows) of the matrix intact. This is the reason for using unimodular operations while manipulating the matrix.

Theorem 11.2. [**Sch86**, Theorem 4.1] *Each rational matrix A of full row rank can be brought to Hermite normal form by a series of elementary unimodular column operations.*

Proof. Let A be a rational matrix of full row rank. Without loss of generality we may assume that A is integral. Start by doing the following three steps: Pick a column of A with a non-zero first entry and make it the first column of A using operation (1). Then using operation (3), we turn all first entries of all other columns of A to zero. Next, using operation (2) if needed, one can ensure that the $(1,1)$-th entry of the current matrix is positive. It's not clear why, for instance, the second step does what it's supposed to do. To prove this, suppose continuing as above, at some intermediate stage we have the matrix $\begin{bmatrix} B & \mathbf{0} \\ C & D \end{bmatrix}$ where B is lower triangular and has a positive diagonal. (The reasoning below can be applied to the start of the procedure as well, when B is vacuous and $D = A$ which proves that we can start as indicated as above.)

Modify D using elementary column operations so that its first row $(\delta_{11}, \ldots, \delta_{1k})$ is non-negative and so that the sum $\delta_{11} + \cdots + \delta_{1k}$ is as small as possible. (There is a minimum sum since the sum is bounded below by zero and is an integer.) By permuting columns, we may assume that $\delta_{11} \geq \delta_{12} \geq \cdots \geq \delta_{1k}$. Then, since A has full row rank, $\delta_{11} > 0$. If $\delta_{12} > 0$, then subtracting the second column of D from the first column of D we will keep the first row of D non-negative but will decrease the sum of the entries in the first row of D which is contradicts the assumption that the sum was as small as possible. So we conclude that $\delta_{12} = \delta_{13} = \cdots = \delta_{1k} = 0$. Thus we have increased the size of B.

Repeating this procedure, we will end with a matrix $[B\ \mathbf{0}]$ with B lower triangular and with a positive diagonal. This also makes B non-singular. Assume that B is a square matrix of size $d(= \text{rank}(A))$. The only task left is to modify B so that the largest entry in any row of B is on the diagonal. For each $i = 1, \ldots, d$ and $j = 1, \ldots, i-1$, add an integer multiple of the i-th column of B to the j-th column of B so that the (i,j)-th entry of B will be non-negative and less than B_{ii}. We do this last step in the order $(i,j) = (2,1), (3,1), (3,2), (4,1), (4,2), (4,3), \ldots$. □

The **Smith normal form** of a matrix A is obtained by doing both elementary column and row operations on A. It has the form $\begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ where D is a diagonal matrix with positive diagonal entries $\delta_1, \ldots, \delta_k$ such that $\delta_1 | \delta_2 | \cdots | \delta_k$. The Smith normal form of a matrix is unique and the product $\delta_1 \cdot \delta_2 \cdot \cdots \cdot \delta_i$ is the g.c.d. of the subdeterminants of A of order i.

It turns out that every rational matrix of full row rank has a unique Hermite normal form, but we will omit that proof here. So it is typical to refer to *the* Hermite normal form of a matrix. The alternative theorem we are after can be stated and proved right away.

Corollary 11.3. [**Sch86**, Corollary 4.1a] *Let $A\mathbf{x} = \mathbf{b}$ be a rational system. Then either this system has an integer solution $\mathbf{x}$ or there is a rational vector $\mathbf{y}$ such that $\mathbf{y}A$ is integral but $\mathbf{yb}$ is not an integer.*

Proof. As usual we first check that both possibilities cannot co-exist. Suppose $\mathbf{x}$ is an integer solution to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{y}$ such that $\mathbf{y}A$ integral and $\mathbf{yb}$ not integer. Then $\mathbf{yb} = \mathbf{y}A\mathbf{x}$ which is a contradiction since $\mathbf{y}A\mathbf{x}$ is an integer.

So suppose that $A\mathbf{x} = \mathbf{b}$ does not have an integral solution. We need to show that then there exists a $\mathbf{y}$ such that $\mathbf{y}A$ is integral but $\mathbf{yb}$ not an integer. We prove the contrapositive. Suppose whenever $\mathbf{y}A$ is integral, $\mathbf{yb}$ is an integer. Then $A\mathbf{x} = \mathbf{b}$ has some solution (possibly fractional) since otherwise, by the alternative theorem from linear algebra, we will have that there is a $\mathbf{y}$ with $\mathbf{y}A = \mathbf{0}$ and $\mathbf{yb} \neq 0$. By scaling the $\mathbf{y}$ with this property, we can ensure for instance that $\mathbf{y}A = \mathbf{0}$ but $\mathbf{yb} = \frac{1}{2}$. So we may assume that the rows of A are linearly independent. Let $[B\ \mathbf{0}]$ be the Hermite normal form of A. Then there is some unimodular matrix U such that $AU = [B\ \mathbf{0}]$ or in other words, $A = [B\ \mathbf{0}]U^{-1}$. If $\mathbf{x}$ is an integer solution to $A\mathbf{x} = \mathbf{b}$ then $[B\ \mathbf{0}](U^{-1}\mathbf{x}) = \mathbf{b}$ and $U^{-1}\mathbf{x}$ is an integer solution to $[B\ \mathbf{0}]\mathbf{z} = \mathbf{b}$. Conversely, if $[B\ \mathbf{0}]\mathbf{z} = \mathbf{b}$ has an integer solution then so does $A\mathbf{x} = \mathbf{b}$. So we may assume without loss of generality that A is in Hermite normal form $[B\ \mathbf{0}]$. Hence we can replace $A\mathbf{x} = \mathbf{b}$ with the square system $B\mathbf{x} = \mathbf{b}$. Since B is non-singular, this system has the unique solution $\mathbf{x} = B^{-1}\mathbf{b}$ which lifts to a solution of $A\mathbf{x} = \mathbf{b}$ by adding zero components.

To complete the proof we just need to argue that $B^{-1}\mathbf{b}$ is integral. Note that $B^{-1}[B\ \mathbf{0}] = [I\ \mathbf{0}]$. If $\mathbf{y}$ is a row of B^{-1}, then what this proves is that $\mathbf{y}A$ is integral. Hence by our assumption, $\mathbf{yb}$ is an integer. Using all the rows of B^{-1}, we conclude that $B^{-1}\mathbf{b}$ is an integral vector as needed. □

In the above proof we have used the non-trivial fact that the Hermite normal form of a matrix A is of the form AU for some unimodular matrix U. See Corollary 4.3b in [**Sch86**] for a proof of this. We could have avoided the use of this fact by simply noting that both statements in the alternative theorem are unaffected by elementary column operations on A. However, we need this unimodular U later and so we might as well start using it.

EXERCISE 11.4. (**HW**) Let $A \in \mathbb{Z}^{m \times n}$ be a matrix of full row rank. Then prove that the following are equivalent.

(1) the g.c.d. of the subdeterminants of A of order m is one.
(2) $A\mathbf{x} = \mathbf{b}$ has an integral solution $\mathbf{x}$ for each integral vector $\mathbf{b}$.
(3) for each $\mathbf{y}$, if $\mathbf{y}A$ is integral then $\mathbf{y}$ is integral.

Our final goal is to show that the problem of deciding whether a rational linear equation system has an integral solution is both in $\mathcal{NP}$ and $co{-}\mathcal{NP}$. Following the usual program, we first need to argue that if $A\mathbf{x} = \mathbf{b}$ has an integer solution then it has one of size polynomially bounded by the size of $(A, \mathbf{b})$.

THEOREM 11.5. [**Sch86**, Theorem 5.2] *The Hermite normal form* $[B\ \mathbf{0}]$ *of a rational matrix* A *of full row rank has size polynomially bounded by the size of* A. *Moreover, there exists a unimodular matrix* U *with* $AU = [B\ \mathbf{0}]$ *such that the size of* U *is polynomially bounded by the size of* A.

PROOF. We may assume that A is integral since multiplying A by a constant also multiplies the Hermite normal form by the same constant. Let $[B\ \mathbf{0}]$ be the Hermite normal form of A and let b_{ii} be the i-th diagonal entry of B. Then note that the product of $b_{11}, \ldots, b_{jj}$ is the determinant of the principal submatrix of $[B\ \mathbf{0}]$ of order j and that all other determinants of order j from the first j rows of $[B\ \mathbf{0}]$ are zero. Therefore, $\prod_{i=1}^{j} b_{ii}$ is the g.c.d. of the subdeterminants of order j of the first j rows of $[B\ \mathbf{0}]$. Now elementary column operations does not change these g.c.d.'s. Hence $\prod_{i=1}^{j} b_{ii}$ is also the g.c.d. of the subdeterminants of order j of the first j rows of A. This implies that the size of $[B\ \mathbf{0}]$ is polynomially bounded by the size of A.

To prove the second statement first assume, by permuting columns if needed, that $A = [A_1, A_2]$ where A_1 is non-singular. Then consider the square matrix $\begin{bmatrix} A_1 & A_2 \\ \mathbf{0} & I \end{bmatrix}$ and its Hermite normal form $\begin{bmatrix} B & \mathbf{0} \\ B_1 & B_2 \end{bmatrix}$. The sizes of B, B_1, B_2 are all polynomially bounded by the size of A. This implies that the size of the unimodular matrix

$$U = \begin{bmatrix} A_1 & A_2 \\ \mathbf{0} & I \end{bmatrix}^{-1} \begin{bmatrix} B & \mathbf{0} \\ B_1 & B_2 \end{bmatrix}$$

is also polynomially bounded by the size of A and $AU = [B\ \mathbf{0}]$. □

COROLLARY 11.6. [**Sch86**, Corollary 5.2a] *If a rational system* $A\mathbf{x} = \mathbf{b}$ *has an integral solution it has one of size polynomially bounded by the sizes of* A *and* $\mathbf{b}$.

PROOF. Assume A has full row rank and that the Hermite normal form of A is $[B\ \mathbf{0}] = AU$ where U is unimodular of size polynomially bounded by the size of A. Let $\mathbf{x}$ be an integral solution of $A\mathbf{x} = \mathbf{b}$. Then

$$B^{-1}\mathbf{b} = B^{-1}A\mathbf{x} = B^{-1}[B\ \mathbf{0}]U^{-1}\mathbf{x} = [I\ \mathbf{0}](U^{-1}\mathbf{x})$$

is integral since $U^{-1}\mathbf{x}$ is integral and has size polynomially bounded by the sizes of A and $\mathbf{b}$. Now check that $\tilde{\mathbf{x}} := U\begin{pmatrix} B^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix}$ is an integral solution of $A\mathbf{x} = \mathbf{b}$. This solution again has size bounded by a polynomial in the sizes of A and $\mathbf{b}$. $\square$

Corollary 11.7. [**Sch86**, Corollary 5.2b] *The problem of deciding whether a rational system $A\mathbf{x} = \mathbf{b}$ has an integer solution lies in $\mathcal{NP} \cap co - \mathcal{NP}$.*

Proof. Since we can test for linear independence of a collection of rows of A in polynomial time in the size of A, we can assume that A has full row rank. If $A\mathbf{x} = \mathbf{b}$ has an integral solution, then by Corollary 11.6, it has one of size polynomially bounded by the sizes of A and $\mathbf{b}$ and hence the above decision problem is in $\mathcal{NP}$. If $A\mathbf{x} = \mathbf{b}$ does not have an integral solution, then by Corollary 11.3 there is a rational $\mathbf{y}$ such that $\mathbf{y}A$ is integral and $\mathbf{yb}$ not an integer. We just need to find such a $\mathbf{y}$ whose size is polynomially bounded in the sizes of A and $\mathbf{b}$.

Let $[B\ \mathbf{0}]$ be the Hermite normal form of A. Then $AU = [B\ \mathbf{0}]$ which implies that

$$B^{-1}A = B^{-1}[B\ \mathbf{0}]U^{-1} = [I\ \mathbf{0}]U^{-1}$$

is integral. On the other hand, if $B^{-1}\mathbf{b}$ was integral then $U\begin{pmatrix} B^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix}$ would be an integral solution of $A\mathbf{x} = \mathbf{b}$. Therefore, $B^{-1}\mathbf{b}$ is not integral. Now let $\mathbf{y}$ be an appropriate row of B^{-1}. Then $\mathbf{y}$ has size polynomially bounded by the size of A as needed. $\square$

It turns out that just like Gaussian elimination, a matrix can be put into Hermite normal form in polynomial time. Therefore, the problem of finding an integral solution to $A\mathbf{x} = \mathbf{b}$ or deciding there is none can be done in polynomial time in the sizes of A and $\mathbf{b}$. See Chapter 5 in [**Sch86**] for details.

CHAPTER 12

The Chvátal-Gomory Procedure

The main goal of this lecture is to present an algorithm for computing the integer hull P^I of a rational polyhedron P. This allows us to compute examples. Let F be a face of the polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} \leq \mathbf{b}\}$. We say that the row $\mathbf{a}_i$ of A, or equivalently the inequality $\mathbf{a}_i\mathbf{x} \leq b_i$, is active at F if for all $\mathbf{x} \in F$, $\mathbf{a}_i\mathbf{x} = b_i$.

Exercise 12.1. (**HW**) Let $\mathcal{A}_F$ consist of the rows of A that are active at the face F of P and let $\mathcal{N}_P(F)$ be the **normal cone** of P at F; i.e.,

$$\mathcal{N}_P(F) := \{\mathbf{c} \in \mathbb{R}^n \,:\, \mathbf{cx} \geq \mathbf{cy} \quad \forall\ \mathbf{x} \in F,\ \mathbf{y} \in P\}.$$

Prove that $\text{cone}(\mathcal{A}_F) = \mathcal{N}_P(F)$. Note that $\mathcal{N}_P(F)$ is precisely the set of all vectors $\mathbf{c}$ that get maximized at F, or equivalently, all $\mathbf{c}$ such that $F \subseteq \text{face}_{\mathbf{c}}(P)$.

Definition 12.2. The rational system $A\mathbf{x} \leq \mathbf{b}$ is **totally dual integral (TDI)** if for each face F of $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$, the rows of A that are active at F form a Hilbert basis.

Exercise 12.3. (**HW**) Prove that the rows of A form a Hilbert basis if and only if $A\mathbf{x} \leq \mathbf{0}$ is TDI. **Hint**. For the "only if" direction, prove that $A\mathbf{x} \leq \mathbf{b}$ is TDI if and only if for each *minimal* face F of $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$, the rows of A that are active at F form a Hilbert basis.

A TDI system $A\mathbf{x} \leq \mathbf{b}$ is *minimally TDI* if any proper subsystem $A'\mathbf{x} \leq \mathbf{b}'$ that also describes $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$ is not TDI. Note that if a TDI system $A\mathbf{x} \leq \mathbf{b}$ is minimally TDI then the following hold:

(1) every inequality in $\mathbf{a}_i\mathbf{x} \leq b_i$ in $A\mathbf{x} \leq \mathbf{b}$ defines a supporting hyperplane of P since otherwise the subsystem obtained by removing this inequality also cuts out P and since $\mathbf{a}_i$ was not active in any face of P, the removal of this inequality does not affect the Hilbert basis property of all the $\mathcal{A}_F$'s, and
(2) no inequality in $A\mathbf{x} \leq \mathbf{b}$ is a non-negative integral combination of others.

Conversely, the above two properties imply that the TDI system $A\mathbf{x} \leq \mathbf{b}$ is minimally TDI.

Theorem 12.4. [**Sch86**, Theorem 22.6] *For each rational polyhedron P there exists a TDI system $A\mathbf{x} \leq \mathbf{b}$ with A integral and $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$. If P is full-dimensional there exists a unique minimal TDI system $A\mathbf{x} \leq \mathbf{b}$ with A integral and $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$. In either case, if P is an integral polyhedron we can choose $\mathbf{b}$ to be integral.*

Proof. For a minimal face F of P, construct a Hilbert basis $\mathbf{a}_1, \ldots, \mathbf{a}_t$ of the normal cone $\mathcal{N}_P(F)$. Pick an $\mathbf{x}_0$ in F and compute $\beta_i := \mathbf{a}_i\mathbf{x}_0$ for each $\mathbf{a}_i$ in this Hilbert basis. Then the inequalities $\mathbf{a}_i\mathbf{x} \leq \beta_i$, $i = 1, \ldots, t$ are all valid for P. Take as $A\mathbf{x} \leq \mathbf{b}$ the union of all such sets of inequalities as F varies over the minimal faces of P. This is a TDI system by construction that describes P. If P is full-dimensional then each normal cone $\mathcal{N}_P(F)$ is pointed and hence has a unique minimal Hilbert basis by Theorem 10.3.

In either case, if P is integral, then we can choose $\mathbf{x}_0$ in each minimal face F to be integral which makes β_i, and hence $\mathbf{b}$, integral. □

Remark 12.5. Note that Example 10.2 can be modified to show that one cannot find a minimal TDI system for a full-dimensional rational polyhedron in polynomial time.

For a vector $\mathbf{b}$ we let $\lfloor \mathbf{b} \rfloor$ denote the vector obtained by replacing each component of $\mathbf{b}$ with its floor.

Algorithm 12.6. **The Chvátal-Gomory procedure**
Input: A rational polyhedron $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$

Initialize: Set $Q := P$
While Q not integral do:

(1) Replace the inequality system describing Q by a TDI system $U\mathbf{x} \leq \mathbf{u}$ that also describes Q.
(2) Let $Q' := \{\mathbf{x} : U\mathbf{x} \leq \lfloor \mathbf{u} \rfloor\}$.
(3) Set $Q := Q'$.

Output: $P^I = Q$.

The Chvátal-Gomory procedure is due to Chvátal [**Chv73**] and Schrijver [**Sch80**]. It is based on the theory of **cutting planes** introduced by Gomory in the sixties. The inequalities in $U\mathbf{x} \leq \lfloor \mathbf{u} \rfloor$ are cutting planes that cut off fractional vertices of the polyhedron Q. We will prove that the Chvátal-Gomory procedure is finite and that it produces the integer hull P^I when it terminates. First we compute an example.

Example 12.7. (c.f. Example 4.6 and Exercise 7.9)
We will compute the matching polytope $M(K_4)$ starting from the polytope $P(K_4)$ described in **Porta** as follows.

```
[thomas@rosa]more matchingk4.ieq
DIM = 6

VALID
1 0 0 0 0 0

INEQUALITIES_SECTION
x1 >= 0
x2 >= 0
x3 >= 0
x4 >= 0
x5 >= 0
x6 >= 0
x1+x6+x4 <= 1
x1+x5+x2 <= 1
x2+x6+x3 <= 1
x4+x5+x3 <= 1
END

[thomas@rosa lecture2]$ traf -v matchingk4.ieq
[thomas@rosa lecture2]$ more matchingk4.ieq.poi
```

```
DIM = 6

CONV_SECTION
(  1)   0   0   0   0   0   0
(  2)   0   0 1/2 1/2   0 1/2
(  3)   0 1/2 1/2   0 1/2   0
(  4) 1/2   0   0 1/2 1/2   0
(  5) 1/2 1/2   0   0   0 1/2
(  6)   0   0   0   0   0   1
(  7)   0   0   0   0   1   0
(  8)   0   0   0   1   0   0
(  9)   0   0   1   0   0   0
( 10)   0   1   0   0   0   0
( 11)   1   0   0   0   0   0
( 12)   0   0   0   0   1   1
( 13)   0   1   0   1   0   0
( 14)   1   0   1   0   0   0

END

strong validity table :
\ I      |            |
 \ N     |            |
P \ E    |            |
 O \ Q   | 1     6    | #
  I \ S  |            |
   N \   |            |
    T \  |            |
     S \ |            |
---------------------------
1        | ***** *.... :  6
2        | **..* .*.** :  6
3        | *..*. *.*** :  6
4        | .**.. ***.* :  6
5        | ..*** .***. :  6
6        | ***** .*.*. :  7
7        | ****. *.*.* :  7
8        | ***.* **..* :  7
9        | **.** *..** :  7
10       | *.*** *.**. :  7
11       | .**** ***.. :  7
12       | ****. .**** :  8
13       | *.*.* ***** :  8
14       | .*.** ***** :  8
          .................
#        | 11111 19999
         | 00000 0
```

Vertices 2,3,4 and 5 are fractional and need to be cut off. From the strong validity table we see that inequalities 1,2,5,7,9,10 are active at vertex 2. Using **Normaliz** we compute the Hilbert basis of the normal cone at vertex 2:

```
[thomas@rosa]more vert2round1.in
6
6
-1 0 0 0 0 0
0 -1 0 0 0 0
0 0 0 0 -1 0
1 0 0 1 0 1
0 1 1 0 0 1
0 0 1 1 1 0
0

[thomas@rosa]normaliz vert2round1
[thomas@rosa]more vert2round1.out
7 generators of integral closure:
  -1   0   0   0   0   0
   0  -1   0   0   0   0
   0   0   0   0  -1   0
   1   0   0   1   0   1
   0   1   1   0   0   1
   0   0   1   1   1   0
   0   0   1   1   0   1
```

The last element in **vert2round1.out** is new and so we need to add the inequality $x_3 + x_4 + x_6 \leq 3/2$ to the TDI system describing $P(K_4)$. The right-hand side 3/2 is computed by evaluating $x_3 + x_4 + x_6$ at vertex 2 which was $(0, 0, 1/2, 1/2, 0, 1/2)$. After rounding down the right-hand-side we will get $x_3 + x_4 + x_6 \leq \lfloor 3/2 \rfloor = 1$ which will cut off the fractional vertex 2. We repeat the same procedure at vertices 3,4 and 5 to get the new inequalities

$$x_2 + x_3 + x_5 \leq 1, \;\; x_1 + x_4 + x_5 \leq 1, \;\; x_1 + x_2 + x_6 \leq 1.$$

After adding these inequalities we get the polytope:

```
[thomas@rosa]more matchingk4_1.ieq
DIM = 6

VALID
1 0 0 0 0 0

INEQUALITIES_SECTION
x1 >= 0
x2 >= 0
x3 >= 0
x4 >= 0
x5 >= 0
x6 >= 0
x1+x4+x6 <= 1
x1+x2+x5 <= 1
x2+x3+x6 <= 1
x3+x4+x5 <= 1
x3+x4+x6 <= 1
x2+x3+x5 <= 1
x1+x4+x5 <= 1
```

```
x1+x2+x6 <= 1
END
```

We then ask **Porta** for its vertex description to discover that we now have the integer hull $M(K_4)$, of $P(K_4)$.

```
[thomas@rosa]traf -v matchingk4_1.ieq

[thomas@rosa lecture2] more matchingk4_1.ieq.poi
DIM = 6

CONV_SECTION
(  1) 0 0 0 0 0 0
(  2) 0 0 0 0 0 1
(  3) 0 0 0 0 1 0
(  4) 0 0 0 1 0 0
(  5) 0 0 1 0 0 0
(  6) 0 1 0 0 0 0
(  7) 1 0 0 0 0 0
(  8) 0 0 0 0 1 1
(  9) 0 1 0 1 0 0
( 10) 1 0 1 0 0 0

END

strong validity table :
\ I      |                  |
 \ N     |                  |
P \ E    |                  |
 O \ Q   | 1     6     11   | #
  I \ S  |                  |
   N \   |                  |
    T \  |                  |
     S \ |                  |
-------------------------------
1        | ***** *.... .... :  6
2        | ***** .*.*. *..* :  9
3        | ****. *.*.* .**. :  9
4        | ***.* **..* *.*. :  9
5        | **.** *..** **.. :  9
6        | *.*** *.**. .*.* :  9
7        | .**** ***.. ..** :  9
8        | ****. .**** **** : 12
9        | *.*.* ***** **** : 12
10       | .*.** ***** **** : 12
          .....................
#        | 88888 86666 6666
```

In this exercise we obtained the integer hull of $P(K_4)$ after one iteration of the Chvátal-Gomory procedure.

In the above example while computing Q' in the first iteration of the Chvátal-Gomory procedure, why was it sufficient to only add in inequalities coming from the Hilbert bases of outer normal cones at the fractional vertices of the initial Q? Can we modify the Chvátal-Gomory procedure in general to incorporate this short-cut?

EXERCISE 12.8. (**HW**) The **stable set polytope**, STAB(G), of a graph G is the convex hull of the characteristic vectors of stable sets in G. For instance, STAB(K_n) is the standard n-dimensional simplex in $\mathbb{R}^n$. A simple linear relaxation of STAB(G) is the polytope

$$\text{FRAC}(G) := \{\mathbf{x} \in \mathbb{R}^V \,:\, x_v \geq 0 \;\forall\; v \in V,\; x_u + x_v \leq 1 \;\forall\; \{u, v\} \in E\}.$$

(1) Prove that STAB(G) is the integer hull of FRAC(G).
(2) Run the Chvátal-Gomory procedure on FRAC(K_4) to obtain STAB(K_4).

We now prove that the Chvátal-Gomory procedure works.

Suppose H is a rational half-space $\{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{cx} \leq \delta\}$ where $\mathbf{c}$ is a primitive integer vector, then note that the integer hull $H^I = \{\mathbf{x} \in \mathbb{R}^n \,:\, \mathbf{cx} \leq \lfloor \delta \rfloor\}$. If $\mathbf{c}$ is not a primitive integer vector then we only get $H^I \subseteq \{\mathbf{x} \,:\, \mathbf{cx} \leq \lfloor \delta \rfloor\}$.

EXAMPLE 12.9. Consider the half-space $H = \{(x, y) \in \mathbb{R}^2 \,:\, y \leq \frac{1}{2}\}$. Then $H^I = \{(x, y) \in \mathbb{R}^2 \,:\, y \leq \lfloor \frac{1}{2} \rfloor = 0\}$. On the other hand if $G = \{(x, y) \in \mathbb{R}^2 \,:\, 2y \leq 1\}$ then $G^I = H^I \subset \{(x, y) \in \mathbb{R}^2 \,:\, 2y \leq \lfloor 1 \rfloor\}$.

DEFINITION 12.10. The **elementary closure** of a rational polyhedron P is $P^{(1)} := \cap\{H^I \,:\, P \subseteq H\}$ where H is a rational half-space containing P.

Note that the intersection in the definition of $P^{(1)}$ can be restricted to rational half-spaces whose hyperplanes support P. Since $P \subseteq H$, we get that $P^I \subseteq H^I$ and hence $P^I \subseteq P^{(1)}$. Let $P^{(i)}$ denote the elementary closure of $P^{(i-1)}$ where $P^{(0)} := P$. Then

$$P \supseteq P^{(1)} \supseteq P^{(2)} \supseteq P^{(3)} \supseteq \cdots \supseteq P^I.$$

In order to prove the Chvátal-Gomory procedure it suffices to prove that
(1) $P^{(1)}$ is a polyhedron, and that
(2) there exists a natural number t such that $P^I = P^{(t)}$.

Recall that by Theorem 12.4, we may assume that $P = \{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$ where $A\mathbf{x} \leq \mathbf{b}$ is TDI and A is integral. The usual definition of a TDI system is different from the one we state in Definition 12.2. It is motivated by optimization. We state this traditional definition as a theorem without proof. For a proof see Theorem 22.5 in [**Sch86**].

THEOREM 12.11. *A rational system $A\mathbf{x} \leq \mathbf{b}$ is TDI if and only if for each integer vector $\mathbf{c}$ for which the the minimum (= maximum) in the LP-duality equation*

$$\min\{\mathbf{yb} \,:\, \mathbf{y} \geq \mathbf{0}, \mathbf{y}A = \mathbf{c}\} = \max\{\mathbf{cx} \,:\, A\mathbf{x} \leq \mathbf{b}\}$$

is finite, the min problem has an integer optimal solution $\mathbf{y}$.

COROLLARY 12.12. *If $A\mathbf{x} \leq \mathbf{b}$ is TDI and $\mathbf{b}$ is integral then the polyhedron $\{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$ is integral.*

PROOF. If $A\mathbf{x} \leq \mathbf{b}$ is TDI then for each integral vector $\mathbf{c}$ for which the max (= min) in the above LP-duality equation is finite, the max is an integer since $\mathbf{b}$ is an integer and the min problem has an integral optimum making the min an integer. Then by Lemma 7.5 (5), $\{\mathbf{x} \,:\, A\mathbf{x} \leq \mathbf{b}\}$ is integral. $\square$

The following theorem shows that $P^{(1)}$ is again a polyhedron and also that in Definition 12.10, we can get away with finitely many half-spaces H.

THEOREM 12.13. [**Sch86**, Theorem 23.1] *Let* $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ *be a polyhedron with* $A\mathbf{x} \leq \mathbf{b}$ *TDI and* A *integral. Then* $P^{(1)} = \{\mathbf{x} : A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor\}$.

PROOF. If $P = \emptyset$ then clearly $P^I = \emptyset$ and the theorem is true. So assume $P \neq \emptyset$. Note that $P^{(1)} \subseteq \{\mathbf{x} : A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor\}$ since each inequality in $A\mathbf{x} \leq \mathbf{b}$ defines a rational half-space H containing P and the corresponding inequality in $A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor$ contains H^I. So we need to prove the reverse inclusion.

Let $H = \{\mathbf{x} : \mathbf{cx} \leq \delta\}$ be a rational half-space containing P. We may assume that $\mathbf{c}$ is a primitive integer vector so that $H^I = \{\mathbf{x} : \mathbf{cx} \leq \lfloor \delta \rfloor\}$. We have

$$\delta \geq \max\{\mathbf{cx} : A\mathbf{x} \leq \mathbf{b}\} = \min\{\mathbf{yb} : \mathbf{y}A = \mathbf{c}, \mathbf{y} \geq \mathbf{0}\}.$$

Since $A\mathbf{x} \leq \mathbf{b}$ is TDI, and the max and therefore, min is finite, and $\mathbf{c}$ is integral, the min is attained by an integral vector $\mathbf{y}_0$. Suppose $\mathbf{x}$ satisfies $A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor$. Then

$$\mathbf{cx} = \mathbf{y}_0 A\mathbf{x} \leq \mathbf{y}_0 \lfloor \mathbf{b} \rfloor \leq \lfloor \mathbf{y}_0 \mathbf{b} \rfloor \leq \lfloor \delta \rfloor$$

which implies that $\mathbf{x} \in H^I$. Since H was an arbitrary rational half-space containing P, we get that $\{\mathbf{x} : A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor\} \subseteq P^{(1)}$. □

To finish our program, it remains to show that there exists a natural integer t such that $P^{(t)} = P^I$.

LEMMA 12.14. [**Sch86**, §23.1] *If* F *is a face of a rational polyhedron* P *then* $F^{(1)} = P^{(1)} \cap F$.

PROOF. Let $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ with A integral and $A\mathbf{x} \leq \mathbf{b}$ TDI. Let $F = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}, \mathbf{ax} = \beta\}$ be a face of P with $\mathbf{a}$ and β integral. Since $A\mathbf{x} \leq \mathbf{b}$ is TDI, the system $A\mathbf{x} \leq \mathbf{b}, \mathbf{ax} \leq \beta$ is also TDI as all it does is add in the redundant inequality $\mathbf{ax} \leq \beta$ to the original TDI description of P. By the same argument, the system $A\mathbf{x} \leq \mathbf{b}, \mathbf{ax} = \beta$ is also TDI. Then since β is integral we get

$$P^{(1)} \cap F = \{\mathbf{x} : A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor, \mathbf{ax} = \beta\} = \{\mathbf{x} : A\mathbf{x} \leq \lfloor \mathbf{b} \rfloor, \mathbf{ax} \leq \lfloor \beta \rfloor, \mathbf{ax} \geq \lceil \beta \rceil\} = F^{(1)}.$$

We have also shown that if $F^{(1)} \neq \emptyset$ then $F^{(1)} = P^{(1)} \cap \{\mathbf{x} : \mathbf{ax} = \beta\}$ is a face of $P^{(1)}$. □

COROLLARY 12.15. *If* F *is a face of* P *and* t *is a natural number then* $F^{(t)} = P^{(t)} \cap F$.

THEOREM 12.16. [**Sch86**, Theorem 23.2] *For a rational polyhedron* P *there exists a natural number* t *such that* $P^{(t)} = P^I$.

PROOF. Let $P \subseteq \mathbb{R}^n$. The proof is by induction on the dimension d of P. If $P = \emptyset$ (i.e., $d = -1$) then $P^I = \emptyset$ and $\emptyset = P^I = P = P^{(0)}$. If P is a point (i.e., $d = 0$), then either $P^I = P$ or $P^I = \emptyset$. In the former case, $t = 0$ works while in the latter case, $t = 1$ works.

So consider $d > 0$ and assume that the theorem holds for all rational polyhedra of dimension less than d. Let the affine hull of P be $\{\mathbf{x} : U\mathbf{x} = \mathbf{v}\}$. If there are no integer points in this affine space, then clearly $P^I = \emptyset$. By Theorem **??**, there exists a rational vector $\mathbf{y}$ such that $\mathbf{y}U =: \mathbf{c}$ is integral but $\mathbf{yv} = \delta$ is not an integer. If $\mathbf{x}$ satisfies $U\mathbf{x} = \mathbf{v}$ then $\mathbf{cx} = \mathbf{y}U\mathbf{x} = \mathbf{yv} = \delta$ and hence $\mathbf{cx} = \delta$ is a supporting hyperplane of P. Then

$$P^{(1)} \subseteq \{\mathbf{x} : \mathbf{cx} \leq \lfloor \delta \rfloor, \mathbf{cx} \geq \lceil \delta \rceil\} = \emptyset$$

and so $t = 1$ works.

So now assume that $\hat{\mathbf{x}}$ is an integer point in the affine hull of P. Since translation by the integer vector $\hat{\mathbf{x}}$ does not affect the theorem, we can assume that the affine hull of P is $\{\mathbf{x} : U\mathbf{x} = \mathbf{0}\}$. We may also assume that U is integral and has full row rank $n-d$. By Theorem 11.2, there exists a unimodular matrix V such that $UV = [W\,\mathbf{0}]$ where W is non-singular. Since V is unimodular, for each $\mathbf{x}$ integral, there exists a unique $\mathbf{z}$ integral such that $\mathbf{x} = V\mathbf{z}$. Thus the affine hull of P is

$$\{\mathbf{x} : U\mathbf{x} = \mathbf{0}\} \cong \{\mathbf{z} : UV\mathbf{z} = \mathbf{0}\} = \{\mathbf{z} : W\mathbf{z} = \mathbf{0}\} = \{\mathbf{0}\}^{n-d} \times \mathbb{R}^d.$$

This implies that P is a full-dimensional polyhedron in $\mathbb{R}^d$ where $\mathbb{R}^n = \mathbb{R}^{n-d} \times \mathbb{R}^d$. To each hyperplane $H = \{\mathbf{x} : \sum_{j=1}^n c_j x_j = \delta\}$ in $\mathbb{R}^n$ we can associate the hyperplane $H' := \{\mathbf{x} : \sum_{j=n-d+1}^n c_j x_j = \delta\}$ of the form $\mathbb{R}^{n-d} \times (\text{hyperplane in}\,\mathbb{R}^d)$. While applying the Chvátal-Gomory procedure to P, it suffices to restrict to hyperplanes of the form H' since pushing H' until it contains an integer point will imply that H will also contain an integer point. Thus we may assume that $n - d = 0$ and P is a full-dimensional polyhedron in $\mathbb{R}^n$.

Since P^I is a polyhedron, there exists a rational matrix M and a rational vector $\mathbf{d}$ such that $P^I = \{\mathbf{x} : M\mathbf{x} \le \mathbf{d}\}$. We may also assume that M is such that there also exists a rational vector $\mathbf{d}'$ such that $P = \{\mathbf{x} : M\mathbf{x} \le \mathbf{d}'\}$ since we can take the rows of M to be the union of all normals of two sets of sufficiently many inequalities that cut out P and P^I respectively and by choosing the right-hand-side vectors $\mathbf{d}$ and $\mathbf{d}'$ to obtain the two polyhedra. Let $\mathbf{m}\mathbf{x} \le d$ be an inequality from $M\mathbf{x} \le \mathbf{d}$ and $H = \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d\}$. We will argue that $P^{(s)} \subseteq H$ for some s. Since there are only finitely many inequalities in $M\mathbf{x} \le \mathbf{d}$, by taking t to be the largest of all the s's, we will get that $P^{(t)} \subseteq P^I$. However, since $P^I \subseteq P^{(t)}$ it will follow that $P^I = P^{(t)}$.

Suppose $P^{(s)}$ is not contained in $H = \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d\}$ for any s. Let $\mathbf{m}\mathbf{x} \le d'$ be the inequality with normal $\mathbf{m}$ in $M\mathbf{x} \le \mathbf{d}'$. Then $P^{(1)} \subseteq \{\mathbf{x} : \mathbf{m}\mathbf{x} \le \lfloor d' \rfloor\}$. Therefore, there exists an integer d'' and an integer r such that

$$\lfloor d' \rfloor \ge d'' > d, \;\; P^{(s)} \subseteq \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d''\} \text{ and } P^{(s)} \not\subseteq \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d'' - 1\} \;\; \forall s \ge r.$$

Let $F := P^{(r)} \cap \{\mathbf{x} : \mathbf{m}\mathbf{x} = d''\}$. Since $\mathbf{m}\mathbf{x} \le d''$ is a valid inequality for $P^{(r)}$, F is a face of $P^{(r)}$, possibly empty, but of dimension less than $d = n$. Moreover, F does not contain any integral vectors since $P^I \subseteq H = \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d\}$ and $d < d''$. By induction, there exists a natural number u such that $F^{(u)} = \emptyset$. Therefore,

$$\emptyset = F^{(u)} = P^{(r+u)} \cap F = P^{(r+u)} \cap \{\mathbf{x} : \mathbf{m}\mathbf{x} = d''\}.$$

So $P^{(r+u)} \subseteq \{\mathbf{x} : \mathbf{m}\mathbf{x} < d''\}$ and hence $P^{(r+u+1)} \subseteq \{\mathbf{x} : \mathbf{m}\mathbf{x} \le d'' - 1\}$ which contradicts our earlier observation. □

Remark 12.17. The above theorem provides a simple proof of the equivalence of (1) and (6) in Lemma 7.5. If each rational supporting hyperplane of P contains an integral vector then $P = P^{(1)}$ and hence $P = P^I$.

CHAPTER 13

Chvátal Rank

Definition 13.1. The **Chvátal rank** of a rational system $A\mathbf{x} \leq \mathbf{b}$ is the smallest integer t such that $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^I = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^{(t)}$.

If $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$, then the Chvátal rank of $A\mathbf{x} \leq \mathbf{b}$ is, roughly speaking, a measure of complexity of P^I relative to P. Note that since $P^{(1)}$ is independent of the initial description of P and the first round of the Chvátal procedure produces $P^{(1)}$, the second round produces $(P^{(1)})^{(1)}$ etc, the Chvátal rank of P is independent of the description of P.

Example 13.2. Example 12.7 shows that the Chvátal rank of the inequality system used to describe $P(K_4)$ is one.

Recall that in Example 7.9, we saw the relaxation $P(G)$ of the matching polytope $M(G)$ of the undirected graph G where $P(G)$ was cut out by the system

$$\mathbf{x}(e) \geq 0 \quad \forall\ e \in E, \quad \sum_{v \in e} \mathbf{x}(e) \leq 1 \quad \forall\ v \in V.$$

Edmonds [**Edm65**] proved that the inequality system

$$\mathbf{x}(e) \geq 0 \quad \forall\ e \in E, \quad \sum_{v \in e} \mathbf{x}(e) \leq 1 \quad \forall\ v \in V, \quad \sum_{e \subseteq U} \mathbf{x}(e) \leq \frac{1}{2}|U| \quad \forall\ U \subseteq V$$

that also describes $P(G)$, is TDI, and that $M(G)$ is obtained by rounding down the right-hand-sides of this TDI system. Thus the Chvátal rank of

$$\mathbf{x}(e) \geq 0 \quad \forall\ e \in E, \quad \sum_{v \in e} \mathbf{x}(e) \leq 1 \quad \forall\ v \in V.$$

describing $P(G)$ is exactly one.

Exercise 13.3. Compute the Chvátal rank of a one-dimensional polyhedron in $\mathbb{R}^1$.

In contrast to the above exercise, we now show that even when $n = 2$, the Chvátal rank of a system $A\mathbf{x} \leq \mathbf{b}$ may be arbitrarily high. This will prove that there is no bound on Chvátal rank of a polyhedron that is a function of dimension alone.

Example 13.4. [**Sch86**, pp.344] Consider the family of matrices

$$A_j := \begin{pmatrix} -1 & 0 \\ 1 & 2j \\ 1 & -2j \end{pmatrix}$$

and the polygons $P_j := \{\mathbf{x} \in \mathbb{R}^2 : A_j\mathbf{x} \leq (0, 2j, 0)^t\}$. The polygon P_j is the convex hull of the points $(0,1), (0,0)$ and $(j, \frac{1}{2})$ and P_j^I is the line segment joining $(0,1)$ and $(0,0)$.

Exercise 13.5. (1) Check that $P_j^{(1)}$ contains the vector $(j-1, \frac{1}{2})$.

(2) Show by induction that $(j-t, \frac{1}{2})$ lies in $P_j^{(t)}$ for $t < j$ and hence $P^{(t)} \neq P_j^I$ for $t < j$.

This proves that the Chvátal rank of the system $A_j\mathbf{x} \leq (0, 2j, 0)^t$ is at least j.

Despite the above example, it is true that Chvátal rank of an inequality system is bounded above by a function of dimension alone when there are no lattice points satisfying the system. This result is due to Cook, Coullard and Turán [**CCT**].

Theorem 13.6. [**Sch86**, Theorem 23.3] *For each natural number d there exists a number $t(d)$ such that if P is a rational polyhedron of dimension d, with $P^I = \emptyset$, then $P^{(t(d))} = \emptyset$.*

Proof. The proof follows by induction on d. If $d = -1$ (i.e., $P = \emptyset$), then $t(-1) := 0$. If $d = 0$ then $t(0) := 1$. As in the proof of Theorem 12.16, we may assume that P is full-dimensional. Note that we can assume that $t(d)$ is an increasing function of d since adding a positive number to $t(d)$ for any d will also serve the same purpose that $t(d)$ serves.

Now a famous result in the Geometry of Numbers states that if a rational polyhedron P contains no lattice points then it has to be "thin" in some direction $\mathbf{c}$. More precisely, if $P^I = \emptyset$, then there exists a primitive integer vector $\mathbf{c}$ and a function $l(d)$ that depends only on dimension such that

$$\max\{\mathbf{cx} : \mathbf{x} \in P\} - \min\{\mathbf{cx} : \mathbf{x} \in P\} \leq l(d).$$

One can take $l(d) = 2d(d+1)2^{\frac{n(n-1)}{4}}$. Let $\delta := \lfloor \max\{\mathbf{cx} : \mathbf{x} \in P\} \rfloor$. We will first prove that for each $k = 0, 1, \ldots, l(d)+1$ we have

$$P^{(k+1+k\cdot t(d-1))} \subseteq \{\mathbf{x} : \mathbf{cx} \leq \delta - k\}. \tag{10}$$

For $k = 0$, (10) says that $P^{(1)} \subseteq \{\mathbf{x} : \mathbf{cx} \leq \delta\}$ which follows from the definition of $P^{(1)}$. Suppose (10) is true for some k. Now consider the face $F := P^{(k+1+k\cdot t(d-1))} \cap \{\mathbf{x} : \mathbf{cx} = \delta - k\}$. Since the dimension of F is less than d, it follows from our induction hypothesis that $F^{(t(d-1))} = \emptyset$. Therefore,

$$(P^{(k+1+k\cdot t(d-1))})^{(t(d-1))} \cap \{\mathbf{x} : \mathbf{cx} = \delta - k\} = F^{(t(d-1))} = \emptyset$$

and hence, $P^{(k+1+(k+1)\cdot t(d-1))} \subseteq \{\mathbf{x} : \mathbf{cx} < \delta - k\}$. This in turn implies that

$$P^{(k+2+(k+1)\cdot t(d-1))} = (P^{(k+1+(k+1)\cdot t(d-1))})^{(1)} \subseteq \{\mathbf{x} : \mathbf{cx} \leq \delta - k - 1\}$$

which shows that (10) holds for $k+1$.

Now taking $k = l(d)+1$ in (10) we have

$$P^{(l(d)+2+(l(d)+1)\cdot t(d-1))} \subseteq \{\mathbf{x} : \mathbf{cx} \leq \delta - l(d) - 1\}.$$

Since $P \subseteq \{\mathbf{x} : \mathbf{cx} > \delta - l(d) - 1\}$ it follows that if we set $t(d) := l(d)+2+(l(d)+1)\cdot t(d-1)$ then $P^{(t(d))} = \emptyset$. □

We now use Theorem 13.6 to prove the main result of this section which says that for a given rational matrix A, there is in fact a finite upper bound on the Chvátal ranks of all inequality systems $A\mathbf{x} \leq \mathbf{b}$ as $\mathbf{b}$ varies. This will allow us to define the *Chvátal rank* of an integer matrix.

Theorem 13.7. [**Sch86**, Theorem 23.4] *For each rational matrix A there exists a number t such that for each rational $\mathbf{b}$, $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^I = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^{(t)}$.*

Proof. Since all the data is rational we may assume that A and $\mathbf{b}$ are integral. We will also assume that A has n columns. Let Δ be the maximum absolute value of a subdeterminant of A. Our goal will be to show that

$$t := \max\{t(n), n^{2n+2}\Delta^{n+1}(1 + t(n-1)) + 1\}$$

will work for the theorem.

Let $P_{\mathbf{b}} := \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$. If $P_{\mathbf{b}}^I = \emptyset$ then the above t works by Theorem 13.6. So assume that $P_{\mathbf{b}}^I \neq \emptyset$. By Theorem 10.12, there exists an integer matrix M that only depends on A and a $\mathbf{d}$ such that $P_{\mathbf{b}}^I = \{\mathbf{x} : M\mathbf{x} \leq \mathbf{d}\}$. Further all entries of M have absolute value at most $n^{2n}\Delta^n$. Let $\mathbf{mx} \leq \delta$ be an inequality from $M\mathbf{x} \leq \mathbf{d}$. We may assume without loss of generality that $\delta = \max\{\mathbf{mx} : \mathbf{x} \in P_{\mathbf{b}}^I\}$. Let $\delta' := \lfloor \max\{\mathbf{mx} : \mathbf{x} \in P\} \rfloor$. Then by Theorem 17.2 [**Sch86**], $\delta' - \delta \leq ||\mathbf{m}||_1 n\Delta \leq n^{2n+2}\Delta^{n+1}$. Now use induction as in Theorem 13.6 to show that for each $k = 0, 1, \ldots, \delta' - \delta$,

$$P_{\mathbf{b}}^{(k+1+k\cdot t(n-1))} \subseteq \{\mathbf{x} : \mathbf{mx} \leq \delta' - k\}.$$

Hence, by taking $k = \delta' - \delta$, we see that $P_{\mathbf{b}}^{(t)} \subseteq \{\mathbf{x} : \mathbf{mx} \leq \delta\}$. As $\mathbf{mx} \leq \delta$ was an arbitrary inequality in $M\mathbf{x} \leq \mathbf{d}$, it follows that $P_{\mathbf{b}}^{(t)} = P_{\mathbf{b}}^I$. □

Definition 13.8. The **Chvátal rank** of a rational matrix A is the smallest t such that $\{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^{(t)} = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}^I$ for each integral vector $\mathbf{b}$.

We extend the definition of a unimodular integer matrix slightly as follows.

Definition 13.9. An integral matrix of rank r is **unimodular** if for each submatrix B consisting of r linearly independent columns of A, the gcd of the subdeterminants of B of order r is one.

Note that this definition does not conflict with our earlier definition in Definition 8.1 since there we only consider integer matrices of full row rank.

Exercise 13.10. (**HW**) Prove that the following conditions are equivalent for an integral matrix A.

(1) A is unimodular.
(2) For each integral $\mathbf{b}$, $\{\mathbf{x} : \mathbf{x} \geq \mathbf{0}, A\mathbf{x} = \mathbf{b}\}$ is integral.
(3) For each integral $\mathbf{c}$, the polyhedron $\{\mathbf{y} : \mathbf{y}A \geq \mathbf{c}\}$ is integral.

Corollary 13.11. *An integral matrix A has Chvátal rank zero if and only if A^t is unimodular.*

Characterizations (of matrices) of higher Chvátal rank are unknown. Chvátal ranks of specific systems have been studied quite a lot.

CHAPTER 14

Complexity of Integer Hulls

The integer hull P^I can be much more complicated than the polyhedron P. If $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ where A has m rows, P has at most m facets and at most $\binom{m}{n}$ vertices. Thus for a rational polyhedron, the number of vertices and facets it can have is bounded above by functions in just m and n. The size of the entries in A and $\mathbf{b}$ do not matter. We first show that there is no function in just m and n that will bound the number of vertices and facets of integer hulls.

Example 14.1. [**Rub70**] Let ϕ_k be the k-th *Fibonacci number* and consider the polytope $P_k \subset \mathbb{R}^2$ defined by the inequalities:

$$\phi_{2k}x + \phi_{2k+1}y \leq \phi_{2k+1}^2 - 1, \;\; x, y \geq 0.$$

The integer hull P_k^I is a polygon with $k + 3$ vertices and facets (edges). See [**Jer71**] for another family of examples in $\mathbb{R}^2$ with only two defining constraints.

The number of facets of P^I can be exponentially large relative to the size of the inequality system defining P.

Theorem 14.2. [**Sch86**, Theorem 18.2] *There is no polynomial ϕ such that for each rational polyhedron $P = \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$, the integer hull P^I has at most $\phi(\text{size}(A, \mathbf{b}))$ facets.*

Proof. Let $n \geq 4$ and let A_n be the vertex-edge incidence matrix of the complete graph K_n. Therefore, A_n is a $n \times \binom{n}{2}$ matrix whose columns are all possible $0, 1$-vectors of length n with exactly two ones. Let

$$P_n := \{\mathbf{x} \geq \mathbf{0} : A_n\mathbf{x} \leq \mathbf{1}\}.$$

We saw in Example **??** that P_n^I is the is the matching polytope of K_n. This integer hull has at least $\binom{n}{2} + 2^{n-1}$ facets as each of the following inequalities determines a facet of P_n^I.

(1) $x(e) \geq 0 \;\; \forall \;\; e \in E$
(2) $\sum_{v \in e} x(e) \leq 1 \;\; \forall \;\; v \in V$
(3) $\sum_{e \subseteq U} x(e) \leq \lfloor \frac{1}{2}|U| \rfloor \;\; \forall \;\; U \subseteq V, |U|$ odd, $|U| \geq 3$

In this situation, $\text{size}(A, \mathbf{b}) = n\binom{n}{2} + \binom{n}{2}3 = \binom{n}{2}(n + 3) = \mathcal{O}(n^3)$ and the number of facets, $\binom{n}{2} + 2^{n-1}$, cannot be bounded by a polynomial in n^3. In fact, Edmonds proved that the above list of inequalities are all the facet inequalities of P_n^I. $\square$

Let us look again at the polyhedron P_n in Theorem 14.2. Each inequality in $A_n\mathbf{x} \leq \mathbf{1}$ has size $1 + \text{size}(\text{a row of } A_n) + \text{size}(1) = 1 + (\binom{n}{2} + (n - 1)) + 1 = \mathcal{O}(n^2)$. Thus the facet complexity of P_n is $\mathcal{O}(n^2)$. Now check that the facet complexity of P_n^I, using the fact that the list of inequalities in Theorem 14.2 define all the facets of P_n^I, is also $\mathcal{O}(n^2)$. This is not a coincidence. We will prove that the facet complexity of P^I is bounded above by a polynomial in the facet complexity of P. At first glance, this seems to contradict Theorem 14.2. But what saves the day is that facet complexity does not care about how many inequalities are needed to describe a polyhedron, but only about the maximum size of any one inequality.

THEOREM 14.3. [**Sch86**, Theorem 17.1] *Let* $P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ *where* $A \in \mathbb{Z}^{m\times n}$ *and* $\mathbf{b} \in \mathbb{Z}^m$. *Then*

$$P^I = \mathrm{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_t) + \mathrm{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s)$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y}_1, \ldots, \mathbf{y}_s$ *are integral vectors with all components at most* $(n+1)\Delta$ *in absolute value, where* Δ *is the maximum absolute value of a subdeterminant of* $[A\ \ \mathbf{b}]$.

PROOF. Assume $P^I \neq \emptyset$. By Theorem 7.2, rec.cone(P^I) = rec.cone(P) = cone$(\mathbf{y}_1, \ldots, \mathbf{y}_s)$. We already saw in the proof of Lemma 10.9 that we can take $\mathbf{y}_1, \ldots, \mathbf{y}_s$ to be integer vectors of infinity norm at most Δ.

Also, we saw in Lecture 5 that if $P = \mathrm{conv}(\mathbf{z}_1, \ldots, \mathbf{z}_k) + \mathrm{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s)$, then each $\mathbf{z}_i$ comes from a minimal face of P and hence a component of $\mathbf{z}_i$ is a quotient of a subdeterminant of $[A\ \ \mathbf{b}]$. Since the absolute value of the numerator of this quotient is bounded above by Δ and the denominator is at least one in absolute value, the quotient and therefore, each component of a $\mathbf{z}_i$, is at most Δ in absolute value.

Now consider the set

$$Z := \{\sum \mu_i \mathbf{y}_i : 0 \le \mu_i \le 1,\ i = 1, \ldots, s \ \text{ and at most } \ n \ \text{ of the } \ \mu_i \ \text{ are non-zero}\}.$$

Every integer point in $\mathrm{conv}(\mathbf{z}_1, \ldots, \mathbf{z}_k) + Z$ has all components of absolute value at most $(n+1)\Delta$ by the above discussion. Let $\mathbf{x}_1, \ldots, \mathbf{x}_t$ be all the integer points in $\mathrm{conv}(\mathbf{z}_1, \ldots, \mathbf{z}_k) + Z$. This will finish the proof if we can argue that every minimal face of P^I contains at least one integer point from $\mathrm{conv}(\mathbf{z}_1, \ldots, \mathbf{z}_k) + Z$. Let F be a minimal face of P^I. Then since P^I is an integral polyhedron, F contains an integer vector $\mathbf{x}^*$. Since $\mathbf{x}^* \in P$, we can write

$$\mathbf{x}^* = \lambda_1 \mathbf{z}_1 + \cdots + \lambda_k \mathbf{z}_k + \mu_1 \mathbf{y}_1 + \cdots + \mu_s \mathbf{y}_s$$

for some $\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_s \ge 0$ and $\sum \lambda_i = 1$. Now every vector in a d-dimensional polyhedral cone lies in the subcone spanned by a collection of d generators of the cone. So we may assume that in the above expression for $\mathbf{x}^*$, at most n of the μ_i's are non-zero. Now consider

$$\tilde{\mathbf{x}} := \mathbf{x}^* - (\lfloor \mu_1 \rfloor \mathbf{y}_1 + \cdots + \lfloor \mu_s \rfloor \mathbf{y}_s)$$

Then $\tilde{\mathbf{x}}$ is an integral vector in $\mathrm{conv}(\mathbf{z}_1, \ldots, \mathbf{z}_k) + Z$ and hence in P. Let $\mathbf{c}$ be an integer vector such that $\mathbf{cx}$ is maximized over P^I at F. Then $\mathbf{cy}_i \le 0$ since $\mathbf{c}$ lies in the polar of the recession cone of P^I. This implies that $\sum_{i=1}^s \lfloor \mu_i \rfloor \mathbf{cy}_i \le 0$ and hence

$$\mathbf{c}\tilde{\mathbf{x}} = \mathbf{cx}^* - \sum_{i=1}^s \lfloor \mu_i \rfloor \mathbf{cy}_i \ge \mathbf{cx}^*.$$

However, since $\tilde{\mathbf{x}} \in P^I$ and $\mathbf{cx}$ is maximized over P^I at F, $\tilde{\mathbf{x}} \in F$. □

COROLLARY 14.4. [**Sch86**, Corollary 17.1a] *Let* $P \subseteq \mathbb{R}^n$ *be a rational polyhedron of facet complexity* ϕ. *Then* P^I *has facet complexity at most* $24n^5\phi \le 24\phi^6$.

PROOF. Since the facet complexity of P is ϕ, there exists some rational inequality system $A\mathbf{x} \le \mathbf{b}$ such that $P = \{\mathbf{x} : A\mathbf{x} \le \mathbf{b}\}$ and each inequality $\mathbf{ax} \le \beta$ has size at most ϕ. We break the proof into several steps.

(1) Suppose $\mathbf{a} = (a_1, \ldots, a_n)$ and $a_i = \frac{p_i}{q_i}$ and $\beta = \frac{p_{n+1}}{q_{n+1}}$ where $q_i > 0$. Then

$$\begin{array}{rcl} \phi & \geq & 1 + \text{size}(\mathbf{a}) + \text{size}(\beta) \\ & = & 1 + n + \sum_{i=1}^{n} \text{size}(a_i) + \text{size}(\beta) \\ & = & 1 + n + n + 1 + \sum_{i=1}^{n+1} \lceil \log_2 |p_i| + 1 \rceil + \sum_{i=1}^{n+1} \lceil \log_2 |q_i| + 1 \rceil \\ & > & \sum_{i=1}^{n+1} \lceil \log_2 |q_i| + 1 \rceil \\ & > & \sum_{i=1}^{n+1} \log_2 q_i \end{array}$$

and hence, $\prod_{i=1}^{n+1} q_i < 2^{\phi}$ which implies that the size of the product of the denominators is at most ϕ.

(2) Clearing denominators in $\mathbf{ax} \leq \beta$ would therefore result in an integral inequality of size at most $(n+2)\phi$. The size of $\mathbf{ax} \leq \beta$ after clearing denominators (by q) is

$$1 + n + \sum_{i=1}^{n} \text{size}(qa_i) + \text{size}(q\beta) \leq \text{size}(\mathbf{ax} \leq \beta) + (n+1)\text{size}(q) \leq (n+2)\phi.$$

Let $\tilde{A}\mathbf{x} \leq \tilde{\mathbf{b}}$ be the resulting integral inequality system describing P.

(3) Since each inequality in $\tilde{A}\mathbf{x} \leq \tilde{\mathbf{b}}$ has size at most $(n+2)\phi$, the size of a square submatrix of $[\tilde{A}\ \tilde{\mathbf{b}}]$ is at most $(n+1)(n+2)\phi \leq (n+2)^2\phi$. Therefore, if Δ is the largest absolute value of a subdeterminant of $[\tilde{A}\ \tilde{\mathbf{b}}]$, then Δ has size at most $2(n+2)^2\phi$ and so $\Delta \leq 2^{2(n+2)^2\phi}$.

(4) By Theorem 14.3, there are integral vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y}_1, \ldots, \mathbf{y}_s$ such that

$$P^I = \text{conv}(\mathbf{x}_1, \ldots, \mathbf{x}_t) + \text{cone}(\mathbf{y}_1, \ldots, \mathbf{y}_s)$$

with each component at most $(n+1)\Delta$ in absolute value.

$$\begin{array}{rcl} \text{size}((n+1)\Delta) & = & 1 + \lceil \log_2((n+1)\Delta + 1) \rceil \\ & \leq & 2 + \log_2((n+1)\Delta + 1) \\ & \leq & 3 + \log_2(n+1) + \log_2 \Delta \\ & \leq & 3 + \log_2(n+1) + 2(n+2)^2\phi \end{array}$$

Therefore, the size of any $\mathbf{x}_i$ or $\mathbf{y}_j$ is at most

$$n + n(3 + \log_2(n+1) + 2(n+2)^2\phi) \leq 6n^3\phi.$$

This implies that the vertex complexity of P^I is at most $6n^3\phi$ and so by Theorem 6.7, the facet complexity of P^I is at most $4n^2(6n^3\phi) = 24n^5\phi$. □

Corollary 14.5. [**Sch86**, Corollary 17.1b,d]

(1) *Let P be a rational polyhedron of facet complexity ϕ. If P contains an integer vector, it contains one of size at most $6n^3\phi$.*

(2) *The problem of deciding whether a rational inequality system $A\mathbf{x} \leq \mathbf{b}$ has an integer solution is in $\mathcal{NP}$.*

Proof. The first statement follows from the proof of Corollary 14.4. If P is a cone, it contains the origin. Else, the $\mathbf{x}_i$'s in the proof of Corollary 14.4 are integral points in P. For the second statement, note that the size of such an $\mathbf{x}_i$ is bounded above by a polynomial in the size of (A, b). □

The feasibility of a rational inequality system over the integers is in general $\mathcal{NP}$-hard. The proof of this is quite involved. We merely show that the problem is in $\mathcal{NP}$.

Bibliography

[BG99] W. Bruns and J. Gubeladze, *Normality and covering properties of affine semigroups*, Journal für die reine und angewandte Mathematik **510** (1999), 161–178.

[BK] W. Bruns and R. Koch, *Normaliz, a program to compute normalizations of semigroups*, Available by anonymous ftp from ftp.mathematik.Uni-Osnabrueck.DE/pub/osm/kommalg/software/.

[CCT] W. Cook, C. Coullard, and Gy. Turan, *On the complexity of cutting-plane proofs*.

[Chv73] V. Chvátal, *Edmonds polytopes and a hierarchy of combinatorial problems*, Discrete Mathematics **4** (1973), 305–337.

[CL] T. Christof and A. Loebel, *PORTA*, available at http://www.zib.de/Optimization/Software/Porta/.

[Edm65] J. Edmonds, *Maximum matching and a polyhedron with 0,1-vertices*, Journal of Research of the National Bureau of Standards (B) **69** (1965), 125–130.

[FZ99] R. Firla and G. Ziegler, *Hilbert bases, unimodular triangulations, and binary covers of rational polyhedral cones*, Discrete and Computational Geometry **21** (1999), 205–216.

[Jer71] R.G. Jeroslow, *Comments on integer hulls of two linear constraints*, Operations Research **19** (1971), 1061–1069.

[Rub70] D.S. Rubin, *On the unlimited number of faces in integer hulls of linear programs with a single constraint*, Operations Research **18** (1970), 940–946.

[Sch80] A. Schrijver, *On cutting planes*, Combinatorics 79 Part II (M. Deza and I.G.Rosenberg, eds.), vol. 9, 1980, pp. 291–296.

[Sch86] ———, *Theory of linear and integer programming*, Wiley-Interscience Series in Discrete Mathematics and Optimization, New York, 1986.

[Tro92] L. E. Trotter, *Lectures on integer programming*, 1992.

Index