4
Integrality of Polyhedra

In this chapter we study properties of polyhedra P which ensure that the Linear Program $\max\left\{c^T x : x \in P\right\}$ has optimal integral solutions.

Definition 4.1 (Integral Polyhedron)
*A polyhedron P is called **integral** if every nonempty face of P contains an integral point.*

Informally speaking, if we are optimizing over an integral polyhedron we get integrality for free: the set of optimal solutions of $z = \max\left\{c^T x : x \in P\right\}$ is a face $F = \left\{x \in P : c^T x = z\right\}$ of P, and, if each face contains an integral point, then there is also an optimal solution which is also integral. In other words, for integral polyhedra we have

$$\max\left\{c^T x : x \in P\right\} = \max\left\{c^T x : x \in P \cap \mathbb{Z}^n\right\}. \tag{4.1}$$

Thus, the IP on the right hand side of (4.1) can be solved by solving the Linear Program on the left hand side of (4.1).

A large part of the study of polyhedral methods for combinatorial optimization problems was motivated by a theorem of Edmonds on matchings in graphs. A matching in an undirected graph $G = (V, E)$ is a set $M \subseteq E$ of edges such that none of the edges in M share a common endpoint. Given a matching M we say that a vertex $v \in V$ is M-*covered* if some edge in M is incident with v. Otherwise, we call v M-*exposed*. Observe that the number of M-exposed nodes is precisely $|V| - 2|M|$. We define:

$$\mathrm{PM}(G) := \left\{\chi^M \in \mathbb{B}^E : M \text{ is a perfect matching in } G\right\} \tag{4.2}$$

to be the set of incidence vectors of perfect matchings in G.

We will show in the next section that a polyhedron P is integral if and only if $P = \mathrm{conv}(P \cap \mathbb{Z}^n)$. Edmonds' Theorem can be stated as follows:

Theorem 4.2 (Perfect Matching Polytope Theorem) *For any graph $G = (V, E)$, the convex hull $\mathrm{conv}(\mathrm{PM}(G))$ of the perfect matchings in G (i.e. each vertex $v \in V$ is incident to exactly one edge of the matching) is identical to the set of solutions of the following linear system:*

$$x(\delta(v)) = 1 \qquad \textit{for all } v \in V \tag{4.3a}$$

$$x(\delta(S)) \geqslant 1 \qquad \textit{for all } S \subseteq V,\ |S| \geqslant 3 \textit{ odd} \tag{4.3b}$$

$$x_e \geqslant 0 \qquad \textit{for all } e \in E. \tag{4.3c}$$

Proof: See Theorem 4.23. □

Observe that any integral solution of (4.3) is a perfect matching. Thus, if P denotes the polyhedron defined by (4.3), then by the equivalence shown in the next section the Perfect Matching Polytope Theorem states that P is integral and $P = \text{conv}(\text{PM}(G))$. Edmond's results is very strong, since it gives us an explicit description of $\text{conv}(\text{PM}(G))$.

4.1 Equivalent Definitions of Integrality

We are now going to give some equivalent definitions of integrality which will turn out to be quite useful later.

Theorem 4.3 *Let $P = P(A, b)$ be a pointed rational polyhedron. Then, the following statements are equivalent:*

(i) P *is an integral polyhedron.*

(ii) The LP $\max\left\{c^T x : x \in P\right\}$ *has an optimal integral solution for all* $c \in \mathbb{R}^n$ *where the value is finite.*

(iii) The LP $\max\left\{c^T x : x \in P\right\}$ *has an optimal integral solution for all* $c \in \mathbb{Z}^n$ *where the value is finite.*

(iv) The value $z^{LP} = \max\left\{c^T x : x \in P\right\}$ *is integral for all* $c \in \mathbb{Z}^n$ *where the value is finite.*

(v) $P = \text{conv}(P \cap \mathbb{Z}^n)$.

Proof: We first show the equivalence of statements (i)-(iv):

(i)⇒(ii) The set of optimal solutions of the LP is a face of P. Since every face contains an integral point, there is an integral optimal solution.

(ii)⇒(iii) trivial.

(iii)⇒(iv) trivial.

(iv)⇒(i) Suppose that (i) is false and let x^0 be an extreme point which by assumption is not integral, say component x^0_j is fractional. By Theorem 3.37 there exists a vector $c \in \mathbb{Z}^n$ such that x^0 is the unique optimal solution of $\max\left\{c^T x : x \in P\right\}$. Since x^0 is the unique optimal solution, we can find a large $\omega \in \mathbb{N}$ such that x^0 is also optimal for the objective vector $\bar{c} := c + \frac{1}{\omega} e_j$, where e_j is the jth unit vector. Clearly, x^0 must then also be optimal for the objective vector $\tilde{c} := \omega\bar{c} = \omega c + e_j$. Now we have

$$\tilde{c}^T x^0 - \omega c^T x^0 = (\omega c^T x^0 + e_j^T x^0) - \omega c^T x^0 = e_j^T x^0 = x^0_j.$$

Hence, at least one of the two values $\tilde{c}^T x^0$ and $c^T x^0$ must be fractional, which contradicts (iv).

We complete the proof of the theorem by showing two implications:

(i)⇒(v) Since P is convex, we have $\text{conv}(P \cap \mathbb{Z}^n) \subseteq P$. Thus, the claim follows if we can show that $P \subseteq \text{conv}(P \cap \mathbb{Z}^n)$. Let $v \in P$, then $v = \sum_{k \in K} \lambda_k x^k + \sum_{j \in J} \mu_j r^j$, where the x^k are the extreme points of P and the r^j are the extreme rays of P. By (i) every x^k is integral, thus $\sum_{k \in K} \lambda_k x^k \in \text{conv}(P \cap \mathbb{Z}^n)$. Since by Observation 3.67 the extreme rays of P and $\text{conv}(P \cap \mathbb{Z}^n)$ are the same, we get that $v \in \text{conv}(P \cap \mathbb{Z}^n)$.

(v)⇒(iv) Let $c \in \mathbb{Z}^n$ be an integral vector. Since by assumption $\operatorname{conv}(P \cap \mathbb{Z}^n) = P$, the LP $\max\{c^T x : x \in P\}$ has an optimal solution in $P \cap \mathbb{Z}^n$ (If $x = \sum_i x^i \in \operatorname{conv}(P \cap \mathbb{Z}^n)$ is a convex combination of points in $P \cap \mathbb{Z}^n$, then $c^T x \leqslant \max_i c^T x^i$ (cf. Observation 2.2)). Thus, the LP has an integral value for every integral $c \in \mathbb{Z}^n$ where the value is finite.

This shows the theorem. $\square$

Recall that each minimal nonempty face of $P(A, b)$ is an extreme point if and only if $\operatorname{rank}(A) = n$ (Corollary 3.27). Thus, we have the following result:

Observation 4.4 *A nonempty polyhedron* $P = P(A, b)$ *with* $\operatorname{rank}(A) = n$ *is integral if and only if all of its extreme points are integral.* $\square$

Moreover, if $P(A, b) \subseteq \mathbb{R}^n_+$ is nonempty, then $\operatorname{rank}(A) = n$. Hence, we also have the following corollary:

Corollary 4.5 *A nonempty polyhedron* $P \subseteq \mathbb{R}^n_+$ *is integral if and only if all of its extreme points are integral.* $\square$

4.2 Matchings and Integral Polyhedra I

As mentioned before, a lot of the interest about integral polyhedra and their applications in combinatorial optimization was fueled by results on the matching polytope. As a warmup we are going to prove a weaker form of the perfect matching polytope theorem due to Birkhoff.

A graph $G = (V, E)$ is called *bipartite*, if there is a partition $V = A \cup B$, $A \cap B = \varnothing$ of the vertex set such that every edge e is of the form $e = (a, b)$ with $a \in A$ and $b \in B$.

Lemma 4.6 *A graph* $G = (V, E)$ *is bipartite if and only if it does not contain an odd cycle.*

Proof: Let $G = (V, E)$ be bipartite with bipartition $V = A \cup B$. Assume for the sake of a contradiction that $C = (v_1, v_2, \ldots, v_{2k-1}, v_{2k} = v_1)$ is an odd cycle in G. We can assume that $v_1 \in A$. Then $(v_1, v_2) \in E$ implies that $v_2 \in B$. Now $(v_2, v_3) \in E$ implies $v_3 \in A$. Continuing we get that $v_{2i-1} \in A$ and $v_{2i} \in B$ for $i = 1, 2, \ldots$. But since $v_1 = v_{2k}$ we have $v_1 \in A \cap B = \varnothing$, which is a contradiction.

Assume conversely that $G = (V, E)$ does not contain an odd cycle. Since it suffices to show that any connected component of G is bipartite, we can assume without loss of generality that G is connected.

Choose $r \in V$ arbitrary. Since G is connected, the shortest path distances from v to all $v \in V$ are finite. We let

$$\begin{aligned} A &= \{v \in V : d(v) \text{ is even}\} \\ B &= \{v \in V : d(v) \text{ is odd}\} \end{aligned}$$

This gives us a partition of V with $r \in A$. We claim that all edges are between A and B. Let $(u, v) \in E$ and suppose that $u, v \in A$. Clearly, $|d(u) - d(v)| \leqslant 1$ which gives us that $d(u) = d(v) = 2k$. Let $p = r, v_1, \ldots, v_{2k} = v$ and $q = r, u_1, \ldots, u_{2k} = u$ be shortest paths from r to v and u, respectively. The paths might share some common parts. Let v_i and u_j be maximal with the property that $v_i = u_j$ and the paths $v_{i+1}, \ldots, v$ and $u_{j+1}, \ldots, u$ are node disjoint. Observe that we must have that $i = j$ since otherwise one of the paths could not be shortest. But then $v_i, v_{i+1}, \ldots, v_{2k} = v, u = u_{2k}, u_{2k-1}, \ldots, u_i = v_i$ is a cycle of odd length, which is a contradiction. $\square$

Theorem 4.7 (Birkhoff's Theorem) *Let G be a bipartite graph. Then,* $\mathrm{conv}(\mathrm{PM}(G)) = P$*, where P is the polytope described by the following linear system:*

$$\begin{aligned} x(\delta(v)) &= 1 && \textit{for all } v \in V && (4.4a) \\ x_e &\geqslant 0 && \textit{for all } e \in E. && (4.4b) \end{aligned}$$

In particular, P is integral.

Proof: Clearly $\mathrm{conv}(\mathrm{PM}(G)) \subseteq P$. To show that $\mathrm{conv}(\mathrm{PM}(G)) = P$ let x be any extreme point of P. Assume for the sake of a contradiction that x is fractional. Define $\tilde{E} := \{e \in E : 0 < x_e < 1\}$ to be set of "fractional edges". Since $x(\delta(v)) = 1$ for any $v \in V$, we can conclude that any vertex that has an edge from $\tilde{E}$ incident with it, in fact is incident to at least two such edges from $\tilde{E}$. Thus, $\tilde{E}$ contains an even cycle C (by Lemma 4.6 the graph G does not contain any odd cycle). Let y be a vector which is alternatingly ± 1 for the edges in C and zero for all other edges. For small $\varepsilon > 0$ we have $x \pm \varepsilon y \in P$. But then, x can not be an extreme point. □

Observe that we can view the assignment problem (see Example 1.6) as the problem of finding a minimum cost perfect matching on a complete bipartite graph. Thus, Birkhoff's theorem shows that we can solve the assignment problem by solving a Linear Program.

Remark 4.8 The concept of total unimodularity derived in the next section will enable us to give an alternative proof of Birkhoff's Theorem.

4.3 Total Unimodularity

Proving that a given polyhedron is integral is usually a difficult task. In this section we derive some conditions under which the polyhedron

$$P^{=}(A, b) = \{x : Ax = b, x \geqslant 0\}$$

is integral for every integral right hand side b.

As a motivation for the following definition of total unimodularity, consider the Linear Program

$$(\mathrm{LP}) \max \left\{ c^T x : Ax = b, x \geqslant 0 \right\}, \tag{4.5}$$

where $\mathrm{rank}\, A = m$. From Linear Programming theory, we know that if (4.5) has a feasible (optimal) solution, it also has a feasible (optimal) basic solution, that is, a solution of the form $x = (x_B, x_N)$, where $x_B = A_{\cdot,B}^{-1} b$ and $x_N = 0$ and $A_{\cdot,B}$ is an $m \times m$ nonsingular submatrix of A indexed by the columns in $B \subseteq \{1, \dots, n\}$, $|B| = m$. Here, $N = \{1, \dots, n\} \setminus B$.

Given such a basic solution $x = (x_B, x_N)$ we have by Cramer's rule:

$$x_i = \frac{\det(B_i)}{\det(A_B)} \quad \text{for } i \in B,$$

where B_i is the matrix obtained from A_B by replacing the ith column by the vector b. Hence, we conclude that if $\det(A_B) = \pm 1$, then each entry of x_B will be integral (provided b is integral as well).

Definition 4.9 (Unimodular matrix, totally unimodular matrix)
*Let A be an $m \times n$-matrix with full row rank. The matrix A is called **unimodular** if all entries of A are integral and each nonsingular $m \times m$-submatrix of A has determinant ± 1. The matrix A is called **totally unimodular**, if each square submatrix of A has determinant ± 1 or 0.*

Since every entry of a matrix forms itself a square submatrix, it follows that for a totally unimodular matrix A every entry must be either ± 1 or 0.

Observation 4.10 *(i) A is totally unimodular, if and only if A^T is totally unimodular.*

(ii) A is totally unimodular, if and only if (A, I) is unimodular.

(iii) A is totally unimodular, if and only if $\begin{pmatrix} A \\ -A \\ I \\ -I \end{pmatrix}$ *is totally unimodular.*

We now show that a Linear Program with a (totally) unimodular matrix always has an integral optimal solution provided the optimum is finite. Thus, by Theorem 4.3 we get that the corresponding polyhedron must be integral.

Theorem 4.11 *Let A be an $m \times n$ matrix with integer entries and linearly independent rows. The polyhedron $\{x \in \mathbb{R}^n : Ax = b, x \geqslant 0\}$ is integral for all $b \in \mathbb{Z}^m$ if and only if A is unimodular.*

Proof: Suppose that A is unimodular and $b \in \mathbb{Z}^m$ is an integral vector. By Corollary 4.5 it suffices to show that all extreme points of $\{x : Ax = b, x \geqslant 0\}$ are integral. Let $\bar{x}$ be such an extreme point. Since A has full row rank, there exists a basis $B \subseteq \{1, \dots, n\}$, $|B| = m$ such that $\bar{x}_B = A_{\cdot,B}^{-1} b$ and $\bar{x}_N = 0$. Since A is unimodular, we have $\det(A_{\cdot,B}) = \pm 1$ and by Cramer's rule we can conclude that $\bar{x}$ is integral.

Assume conversely that $\{x : Ax = b, x \geqslant 0\}$ is integral for every integral vector b. Let B be a basis of A. We must show that $\det(A_{\cdot,B}) = \pm 1$. Let $\bar{x}$ be the extreme point corresponding to the basis B. By assumption $\bar{x}_B = A_{\cdot,B}^{-1} b$ is integral for all integral b such that $\bar{x}_B \geqslant 0$. Consider for $j \in \{1, \dots, m\}$ and some $z \in \mathbb{Z}^n$ the vector $w_j := A_{\cdot,B} z + e_j$, where e_j is the jth unit vector. Then $A_{\cdot,B}^{-1} w_j = z + A_{\cdot,B}^{-1} e_j$. We can choose $z \in \mathbb{Z}^n$ such that $A_{\cdot,B}^{-1} w_j \geqslant 0$ for all $j = 1, \dots, m$. Thus from $z \in \mathbb{Z}^n$ and $A_{\cdot,B}^{-1} w_j \in \mathbb{Z}^n$ we can conclude that $A_{\cdot,B}^{-1} e_j$ is integral for all j, which gives that $A_{\cdot,B}^{-1}$ must be integral. Thus, it follows that $\det(A_{\cdot,B}^{-1}) = 1/\det(A_{\cdot,B})$ is integral. On the other hand, $\det(A_{\cdot,B})$ is also integral by the integrality of A. Hence, $\det(A_{\cdot,B}) = \pm 1$ as required. $\square$

We use the result of the previous theorem to show the corresponding result for the polyhedron $\{x : Ax \leqslant b, x \geqslant 0\}$.

Corollary 4.12 (Integrality-Theorem of Hoffmann and Kruskal) *Let A be an $m \times n$ matrix with integer entries. The matrix A is totally unimodular if and only if the polyhedron $P(b) := \{x : Ax \leqslant b, x \geqslant 0\}$ is integral for all $b \in \mathbb{Z}^m$.*

Proof: The proof is very simular to Theorem 4.11. It is easy to see that x is an extreme point of $P(b)$ if and only if (x, y) is an extreme point of

$$\{(x, y) : Ax + Iy = b, x, y, \geqslant 0\},$$

where y is uniquely determined by $y = b - Ax$. Thus, if A is totally unimodular, then (A, I) is (totally) unimodular and integrality of $P(b)$ follows directly from Theorem 4.11.

Assume now conversely that $P(b)$ is integral for each integral b. Let A_1 be a $k \times k$-submatrix of A and let

$$\bar{A} := \begin{pmatrix} A_1 & 0 \\ A_2 & I_{m-k} \end{pmatrix}$$

be the $m \times m$-submatrix of (A, I) generated from A_1 by taking the appropriate $m-k$ unit vectors from I. Notice that $\bar{A}$ is nonsingular if and only if A_1 is. Consider the vectors $w_j := \bar{A}z + e_j$, $j = 1, \dots, m$. Then $\bar{A}^{-1}w_j = z + \bar{A}^{-1}e_j$. We can choose $z \in \mathbb{Z}^n$ such that $\bar{A}^{-1}w_j \geqslant 0$ for all $j = 1, \dots, m$. Hence, $z + \bar{A}^{-1}e_j$ is the vector of basis variables of an extreme point of $P(b)$ which by assumption is integral. Again, we can conclude that $\bar{A}^{-1}$ is integral and so must be A_1^{-1}. By the same argument as in Theorem 4.11 we get that $\det(A_1) \in \{+1, -1\}$. □

The Integrality-Theorem of Hoffmann and Kruskal in conjunction with Observation 4.10 yields more characterizations of totally unimodular matrices.

Corollary 4.13 *Let A be an integral matrix. Then the following statements hold:*

(a) A is totally unimodular, if and only if the polyhedron $\{x : a \leqslant Ax \leqslant b, l \leqslant x \leqslant u\}$ is integral for all integral a, b, l, u.

(b) A is totally unimodular, if and only if the polyhedron $\{x : Ax = b, 0 \leqslant x \leqslant u\}$ is integral for all integral b, u.

□

4.4 Conditions for Total Unimodularity

In this section we derive sufficient conditions for a matrix to be totally unimodular.

Theorem 4.14 *Let A be any $m \times n$ matrix with entries taken from $\{0, +1, -1\}$ with the property that any column contains at most two nonzero entries. Suppose also that there exists a partition $M_1 \cup M_2 = \{1, \dots, m\}$ of the rows of A such that every column j with two nonzero entries satisfies: $\sum_{i \in M_1} a_{ij} = \sum_{i \in M_2} a_{ij}$. Then, A is totally unimodular.*

Proof: Suppose for the sake of a contradiction that A is not totally unimodular. Let B be a smallest square submatrix such that $\det(B) \notin \{0, +1, -1\}$. Obviously, B can not contain any column with at most one nonzero entry, since otherwise B would not be smallest. Thus, any column of B contains exactly two nonzero entries. By the assumptions of the theorem, adding the rows in B that are in M_1 and subtracting those that are in M_2 gives the zero vector, thus $\det(B) = 0$, a contradiction! □

Example 4.15
Consider the LP-relaxation of the assignment problem.

$$\min \quad \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{4.6a}$$

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad \text{for } j = 1, \dots, n \tag{4.6b}$$

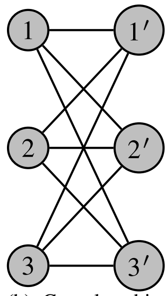
$$\sum_{j=1}^{n} x_{ij} = 1 \qquad \text{for } i = 1, \dots, n \tag{4.6c}$$

$$0 \leqslant x \leqslant 1, \tag{4.6d}$$

We can write the constraints (4.6b) and (4.6c) as $Ax = 1$, where A is the node-edge incidence matrix of the complete bipartite graph $G = (V, E)$ (Figure 4.1 shows the situation for $n = 3$). The rows of A correspond to the vertices and the columns to the edges. The column corresponding to edge (u, v) has exactly two ones, one at the row for u and one at the row for v. The fact that G is bipartite $V = A \cup B$, gives us a partition $A \cup B$ of the rows such that the conditions of Theorem 4.14 are satisfied. Hence, A is totally unimodular.

$$\left(\begin{array}{ccccccccc} 1 & 1 & 1 & & & & & & \\ & & & 1 & 1 & 1 & & & \\ & & & & & & 1 & 1 & 1 \\ \hline 1 & & & 1 & & & 1 & & \\ & 1 & & & 1 & & & 1 & \\ & & 1 & & & 1 & & & 1 \end{array}\right) \cdot \begin{pmatrix} x_{11'} \\ x_{12'} \\ x_{13'} \\ x_{21'} \\ x_{22'} \\ x_{23'} \\ x_{31'} \\ x_{32'} \\ x_{33'} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

(a) Constraint system



(b) Complete bipartite graph

Figure 4.1: The matrix of the assignment problem as the node-edge incidence matrix of a complete bipartite graph.

$\triangleleft$

We derive some other useful consequences of Theorem 4.14:

Theorem 4.16 *Let* A *be any* $m \times n$ *matrix with entries taken from* $\{0, +1, -1\}$ *with the property that any column contains at most one* $+1$ *and at most one* -1. *Then* A *is totally unimodular.*

Proof: First, assume that A contains *exactly* two nonzero entries per column. The fact that A is totally unimodular for this case follows from Theorem 4.14 with $M_1 = \{1, \ldots, m\}$ and $M_2 = \emptyset$. For the general case, observe that a column with at most one nonzero from $\{-1, +1\}$ can not destroy unimodularity, since we can develop the determinant (of a square submatrix) by that column. $\square$

The node-arc incidence matrix of a directed network $G = (V, A)$ is the $n \times m$-Matrix $M(A) = (m_{xy})$ such that

$$m_{xa} = \begin{cases} +1 & \text{if } a = (i,j) \text{ and } x = j \\ -1 & \text{if } a = (i,j) \text{ and } x = i \\ 0 & \text{otherwise} \end{cases}$$

The minimum cost flow problem can be stated as the following Linear Program:

$$\min \sum_{(i,j) \in A} c(i,j)x(i,j) \tag{4.7a}$$

$$\sum_{j:(j,i) \in A} x(j,i) - \sum_{j:(i,j) \in A} x(i,j) = b(i) \qquad \text{for all } i \in V \tag{4.7b}$$

$$0 \leqslant x(i,j) \leqslant u(i,j) \qquad \text{for all } (i,j) \in A \tag{4.7c}$$

By using the node-arc incidence matrix $M = M(A)$, we can rewrite (4.7) as:

$$\min \left\{ c^T x : Mx = b, 0 \leqslant x \leqslant u \right\}, \tag{4.8}$$

where b is the vector of all required demands.

Corollary 4.17 *The node-arc incidence matrix of a directed network is totally unimodular.*

Proof: The claim follows immediately from Theorem 4.16 and Corollary 4.13. □

We close this section by one more sufficient condition for total unimodularity.

Theorem 4.18 (Consecutive ones Theorem) *Let A be any $m \times n$-matrix with entries from $\{0,1\}$ and the property that the rows of A can be permutated in such a way that all 1s appear consecutively. Then, A is totally unimodular.*

Proof: Let B be a square submatrix of A. Without loss of generality we can assume that the rows of A (and thus also of B) are already permuted in such a way that the ones appear consecutively. Let $b_1^T, \dots, b_k^T$ be the rows of B. Consider the matrix B' with rows $b_1^T - b_2^T, b_2^T - b_3^T, \dots, b_{k-1}^T - b_k^T, b_k$. The determinant of B' is the same as of B.

Any column of B' contains at most two nonzero entries, one of which is a -1 (one before the row where the ones in this column start) and a $+1$ (at the row where the ones in this column end). By Theorem 4.16, B' is totally unimodular, in particular $\det(B') = \det(B) \in \{0, +1, -1\}$. □

4.5 Applications of Unimodularity: Network Flows

We have seen above that if M is the node-arc incidence matrix of a directed graph, then the polyhedron $\{x : Mx = b, 0 \leqslant x \leqslant u\}$ is integral for all integral b and u. In particular, for integral b and u we have strong duality between the IP

$$\max\left\{c^T x : Mx = b, 0 \leqslant x \leqslant u, x \in \mathbb{Z}^n\right\} \tag{4.9}$$

and the dual of the LP-relaxation

$$\min\left\{b^T z + u^T y : M^T z + y \geqslant c, y \geqslant 0\right\}. \tag{4.10}$$

Moreover, if the vector c is integral, then by total unimodularity the LP (4.10) has always an integral optimal solution value.

4.5.1 The Max-Flow-Min-Cut-Theorem

As an application of the strong duality of the problems (4.9) and (4.10) we will establish the Max-Flow-Min-Cut-Theorem.

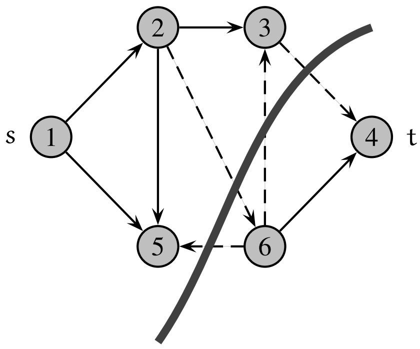
Definition 4.19 (Cut in a directed graph, forward and backward part)
Let $G = (V, A)$ be a directed graph and $S \cup T = V$ a partition of the node set V. We call (S, T) the ***cut induced by*** *S* ***and*** *T. We also denote by*

$$\delta^+(S) := \{(i,j) \in A : i \in S \text{ und } j \in T\}$$
$$\delta^-(S) := \{(j,i) \in A : j \in T \text{ und } i \in S\}$$

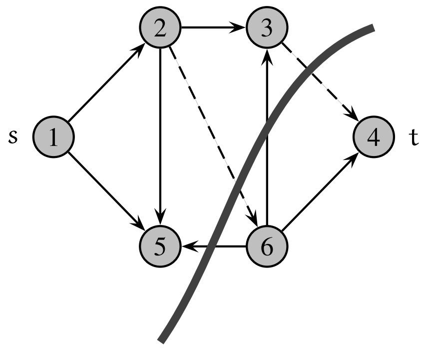
the ***forward part*** *and the* ***backward part*** *of the cut. The cut (S, T) is an (s,t)-cut if $s \in S$ and $t \in T$.*

If $u \colon A \to \mathbb{R}_{\geqslant 0}$ is a capacity function defined on the arcs of the network $G = (V, A)$ and (S, T) is a cut, then the ***capacity of the cut*** *is defined to be the sum of the capacities of its forward part:*

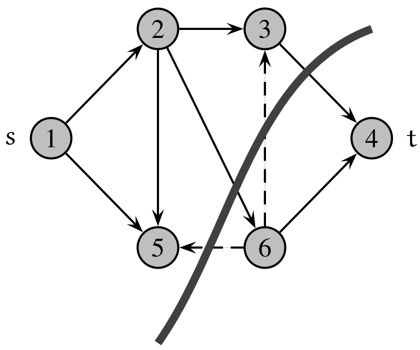
$$u(\delta^+(S)) := \sum_{(u,v) \in (S,T)} u(u,v).$$



(a) An (s,t)-cut (S,T) in a directed graph. The arcs in $\delta^+(S) \cup \delta^-(S)$ are shown as dashed arcs.



(b) The forward part $\delta^+(S)$ of the cut: Arcs in $\delta^+(S)$ are shown as dashed arcs.



(c) The backward part $\delta^-(S)$ of the cut: arcs in $\delta^-(S)$ are shown as dashed arcs.

Figure 4.2: A cut (S,T) in a directed graph and its forward part $\delta^+(S)$ and backward part $\delta^-(S)$.

Figure 4.2 shows an example of a cut and its forward and backward part.

Let f be an (s,t)-flow and (S,T) be an (s,t)-cut in G. For a node $i \in V$ we define by

$$\text{excess}_f(i) := \sum_{a\in\delta^-(v)} f(a) - \sum_{a\in\delta^+(v)} f(a) \tag{4.11}$$

the *excess* of i with respect to f. The first term in (4.11) corresponds to the inflow into i, the second term is the outflow out of i. Then we have:

$$\begin{aligned} \text{val}(f) &= -\text{excess}_f(s) = -\sum_{i\in S} \text{excess}_f(i) \\ &= \sum_{i\in S}\left(\sum_{(i,j)\in A} f(i,j) - \sum_{(j,i)\in A} f(j,i)\right). \end{aligned} \tag{4.12}$$

If for an arc (x,y) both nodes x and y are contained in S, then the term $f(x,y)$ appears twice in the sum (4.12), once with a positive and once with a negative sign. Hence, (4.12) reduces to

$$\text{val}(f) = \sum_{a\in\delta^+(S)} f(a) - \sum_{a\in\delta^-(S)} f(a). \tag{4.13}$$

Using that f is feasible, that is, $0 \leqslant f(i,j) \leqslant u(i,j)$ for all arcs (i,j), we get from (4.13):

$$\text{val}(f) = \sum_{a\in\delta^+(S)} f(a) - \sum_{a\in\delta^-(S)} f(a) \leqslant \sum_{a\in\delta^+(S)} u(a) = u(\delta^+(S)).$$

Thus, the value $\text{val}(f)$ of the flow is bounded from above by the capacity $u(\delta^+(S)))$ of the cut. We have proved the following lemma:

Lemma 4.20 *Let f be an (s,t)-flow and (S,T) an (s,t)-cut. Then:*

$$\mathit{val}(f) \leqslant u(\delta^+(S)).$$

Since f and $[S,T]$ are arbitrary we deduce that:

$$\max_{f \text{ is an } (s,t)\text{-flow in } G} \mathit{val}(f) \leqslant \min_{(S,T) \text{ is an } (s,t)\text{-cut in } G} u(\delta^+(S)). \tag{4.14}$$

□

We are now ready to prove the famous Max-Flow-Min-Cut-Theorem of Ford and Fulkerson:

Theorem 4.21 (Max-Flow-Min-Cut-Theorem) *Let $G = (V,A)$ be a network with capacities $u\colon A \to \mathbb{R}_+$, then the value of a maximum (s,t)-flow equals the minimum capacity of an (s,t)-cut.*

Proof: We add a backward arc (t,s) to G. Call the resulting graph $G' = (V,A')$, where $A' = A \cup \{(t,s)\}$. Then, we can write the maximum flow problem as the Linear Program

$$z = \max\{x_{ts} : Mx = 0, 0 \leqslant x \leqslant u\}, \tag{4.15}$$

where M is the node-arc incidence matrix of G' and $u(t,s) = +\infty$. We know that M is totally unimodular from Corollary 4.17. So, (4.15) has an optimal integral solution value for all integral capacities u. By Linear Programming duality we have:

$$\max\{x_{ts} : Mx = 0, 0 \leqslant x \leqslant u\} = \min\left\{u^T y : M^T z + y \geqslant \chi^{(t,s)}, y \geqslant 0\right\},$$

where $\chi^{(t,s)}$ is the vector in $\mathbb{R}^A$ which has a one at entry (t,s) and zero at all other entries. We unfold the dual which gives:

$$
\begin{aligned}
w = \min & \sum_{(i,j)\in A} u_{ij} y_{ij} && && (4.16a)\\
& z_i - z_j + y_{ij} \geqslant 0 && \text{for all } (i,j) \in A && (4.16b)\\
& z_t - z_s \geqslant 1 && && (4.16c)\\
& y_{ij} \geqslant 0 && \text{for all } (i,j) \in A && (4.16d)
\end{aligned}
$$

There are various ways to see that (4.16) has an optimum solution which is also integral, for instance:

- The constraint matrix of (4.16) is of the form $(M^T I)$ and, from the total unimodularity of M it follows that $(M^T I)$ is also totally unimodular. In particular, (4.16) has an integral optimal solution for every integral right hand side (and our right hand side is integral!).
- The polyhedron of the LP (4.15) is integral by total unimodularity. Thus, it has an optimum integer value for all integral capacities (the objective is also integral). Hence, by LP-duality (4.16) has an optimum integral value for all integral objectives (which are the capacities). Hence, by Theorem 4.3 the polyhedron of (4.16) is integral and (4.16) has an optimum integer solution.

Let (y^*, z^*) be such an integral optimal solution of (4.16). Observe that replacing z^* by $z^* - \alpha$ for some $\alpha \in \mathbb{R}$ does not change anything, so we may assume without loss of generality that $z_s^* = 0$.

Since (y^*, z^*) is integral, the sets S and T defined by

$$
\begin{aligned}
S &:= \{v \in V : z_v^* \leqslant 0\}\\
T &:= \{v \in V : z_v^* \geqslant 1\}
\end{aligned}
$$

induce an (S,T)-cut. Then,

$$
w = \sum_{(i,j)\in A} u_{ij} y_{ij}^* \geqslant \sum_{(i,j)\in\delta^+(S)} u_{ij} y_{ij}^* \geqslant \sum_{(i,j)\in\delta^+(S)} u_{ij} (\underbrace{z_j^*}_{\geqslant 1} - \underbrace{z_i^*}_{\leqslant 0}) \geqslant u(\delta^+(S)).
$$

Thus, the optimum value w of the dual (4.16) which by strong duality equals the maximum flow value is at least the capacity $u(\delta^+(S))$ of the cut (S,T). By Lemma 4.20 it now follows that (S,T) must be a minimum cut and the claim of the theorem is proved. □

4.6 Matchings and Integral Polyhedra II

Birkhoff's theorem provided a complete description of $\mathrm{conv}(\mathrm{PM}(G))$ in the case where the graph $G = (V,E)$ was bipartite. In general, the conditions in (4.4) do not suffice to ensure integrality of every extreme point of the corresponding polytope. Let $\mathrm{FPM}(G)$ (the *fractional perfect matching polytope*) denote the polytope defined by (4.4). Consider the case where the graph G contains an odd cycle of length 3 (cf. Figure 4.3).

The vector $\tilde{x}$ with $\tilde{x}_{e_1} = \tilde{x}_{e_2} = \tilde{x}_{e_3} = \tilde{x}_{e_5} = \tilde{x}_{e_6} = \tilde{x}_{e_7} = 1/2$ and $\tilde{x}_{e_4} = 0$ is contained in $\mathrm{FPM}(G)$. However, $\tilde{x}$ is not a convex combination of incidence vectors of perfect matchings of G, since $\{e_3, e_4, e_5\}$ is the only perfect matching in G. However, the fractional matching polytope $\mathrm{FPM}(G)$ still has an interesting structure, as the following theorem shows:

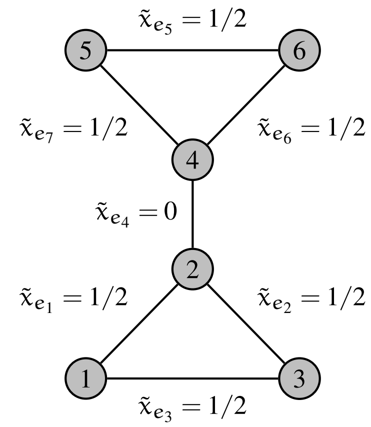


Figure 4.3: In an odd cycle, the blossom inequalities are necessary to ensure integrality of all extreme points.

Theorem 4.22 (Fractional Perfect Matching Polytope Theorem) *Let* $G = (V, E)$ *be a graph and* $x \in \mathrm{FPM}(G)$. *Then,* x *is an extreme point of* $\mathrm{FPM}(G)$ *if and only if* $x_e \in \{0, 1/2, 1\}$ *for all* $e \in E$ *and the edges* e *for which* $x_e = 1/2$ *form node disjoint odd cycles.*

Proof: Suppose that $\tilde{x}$ is a half-integral solution satisfying the conditions stated in the theorem. Define the vector $w \in \mathbb{R}^n$ by $w_e = -1$ if $\tilde{x}_e = 0$ and $w_e = 0$ if $\tilde{x}_e > 0$. Consider the face $F = \{x \in \mathrm{FPM}(G) : w^T x = 0\}$. Clearly, $\tilde{x} \in F$. We claim that $F = \{\tilde{x}\}$ which shows that $\tilde{x}$ is an extreme point.

For every $x \in F$ we have

$$0 = w^T x = -\sum_{e \in E : \tilde{x}_e = 0} \underbrace{x_e}_{\geqslant 0}.$$

Thus $x_e = 0$ for all edges such that $\tilde{x}_e = 0$. Now consider an edge e where $\tilde{x}_e = 1/2$. By assumption, this edge lies on an odd cycle C. It is now easy to see that the values of x on the cycle must be alternatingly θ and $1 - \theta$ since $x(\delta(v)) = 1$ for all $v \in V$ (see Figure 4.4). The only chance that $x \in \mathrm{FPM}(G)$ is $\theta = 1/2$ and thus $x = \tilde{x}$.

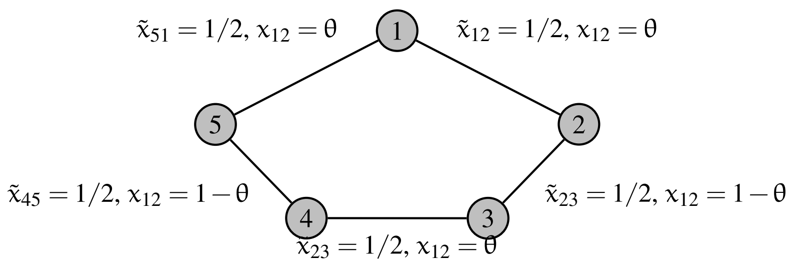
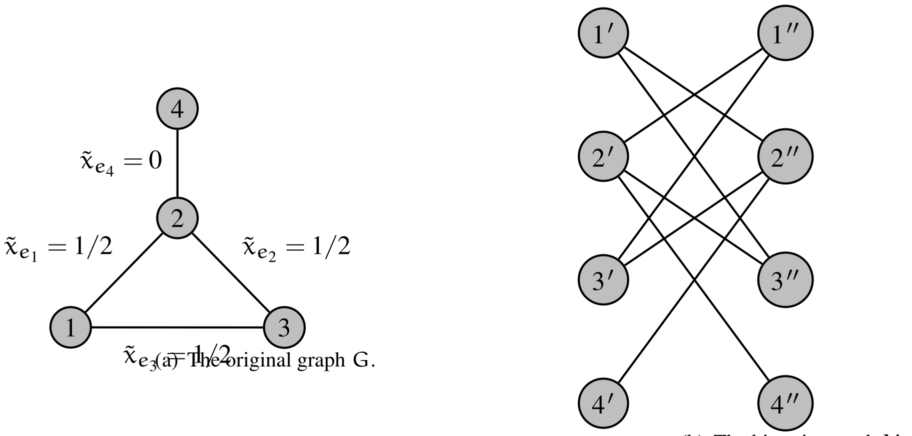


Figure 4.4: If $\tilde{x}$ satisfies the conditions of Theorem 4.22 it is the only member of the face $F = \{x \in \mathrm{FPM}(G) : w^T x = 0\}$.

Assume conversely that $\tilde{x}$ is an extreme point of $\mathrm{FPM}(G)$. We first show that $\tilde{x}$ is half-integral. By Theorem 3.37 there is an integral vector c such that $\tilde{x}$ is the unique solution of $\max\{c^T x : x \in \mathrm{FPM}(G)\}$.

Construct a bipartite graph $H = (V_H, E_H)$ from G by replacing each node $v \in V$ by two nodes v', v'' and replacing each edge $e = (u, v)$ by two edges $e' = (u', v'')$ and $e'' = (v', u'')$ (see Figure 4.5 for an illustration). We extend the weight function $c \colon E \to \mathbb{R}$ to E_H by setting $c(u', v'') = c(v', u'') = c(u, v)$.



(a) The original graph G.

(b) The bipartite graph H.

Figure 4.5: Construction of the bipartite graph H in the proof of Theorem 4.22.

Observe that, if $x \in \text{FPM}(G)$, then x' defined by $x'_{u',v''} := x'_{v',u''} := x_{uv}$ is a vector in $\text{FPM}(H)$ of twice the objective function value of x. Conversely, if $x' \in \text{FPM}(H)$, then $x_{uv} = \frac{1}{2}(x'_{u'v''} + x'_{u''v'})$ is a vector in $\text{FPM}(G)$ of half of the objective function value of x'.

By Birkhoff's Theorem (Theorem 4.7), the problem

$$\max\left\{c^T x_H : x_H \in \text{FPM}(H)\right\}$$

has an integral optimal solution x^*_H. Using the correspondence $x_{uv} = \frac{1}{2}(x^*_{u'v''} + x^*_{u''v'})$ we obtain a half-integral optimal solution to

$$\max\left\{c^T x : x \in \text{FPM}(G)\right\}.$$

Since $\tilde{x}$ was the unique optimal solution to this problem, it follows that $\tilde{x}$ must be half-integral.

If $\tilde{x}$ is half-integral, it follows that the edges $\{e : \tilde{x}_e = 1/2\}$ must form node disjoint cycles (every node that meets a half-integral edge, meets exactly two of them). As in the proof of Birkhoff's Theorem, none of these cycles can be even, since otherwise $\tilde{x}$ is no extreme point. □

With the help of the previous result, we can now prove the Perfect Matching Polytope Theorem, which we restate here for convenience.

Theorem 4.23 (Perfect Matching Polytope Theorem) *For any graph* $G = (V, E)$*, the convex hull* $\text{conv}(\text{PM}(G))$ *of the perfect matchings in* G *is identical to the set of solutions of the following linear system:*

$$x(\delta(v)) = 1 \qquad \textit{for all } v \in V \tag{4.17a}$$

$$x(\delta(S)) \geqslant 1 \qquad \textit{for all } S \subseteq V,\ |S| \geqslant 3 \textit{ odd} \tag{4.17b}$$

$$x_e \geqslant 0 \qquad \textit{for all } e \in E. \tag{4.17c}$$

The inequalities (4.17b) are called blossom inequalities.

Proof: We show the claim by induction on the number $|V|$ of vertices of the graph $G = (V, E)$. If $|V| = 2$, then the claim is trivial. So, assume that $|V| > 2$ and the claim holds for all graphs with fewer vertices.

Let P be the polyhedron defined by the inequalities (4.17) and let $x' \in P$ be any extreme point of P. Since $\text{conv}(\text{PM}(G)) \subseteq P$, the claim of the theorem follows if we can show that $x' \in \text{PM}(G)$. Since $\{x'\}$ is a minimal face of $\text{conv}(\text{PM}(G))$, by Theorem 3.6 there exist a subset $E' \subseteq E$ of the edges and a family $\mathcal{S}'$ of odd subsets $S \subseteq V$ such that x' is the unique solution to:

$$
\begin{aligned}
x(\delta(v)) &= 1 && \text{for all } v \in V && (4.18a)\\
x(\delta(S)) &= 1 && \text{for all } S \in \mathcal{S}' && (4.18b)\\
x_e &= 0 && \text{for all } e \in E'. && (4.18c)
\end{aligned}
$$

Case 1: $\mathcal{S}' = \emptyset$.
In this case, x' is a vertex of $\text{FPM}(G)$. By Theorem 4.22, x' is half-integral and the fractional edges form node-disjoint odd cycles. On the other hand, x' satisfies the blossom inequalities (4.17b) which is a contradiction.

Case 2: $\mathcal{S}' \neq \emptyset$.
Fix $S \in \mathcal{S}'$, by definition we have $x'(\delta(S)) = 1$. Notice that $|S|$ is odd. The complement $\bar{S} := V \setminus S$ need not be of odd cardinality, but observe that, if $\bar{S}$ is of even cardinality, then G does not contain a perfect matching (since the total number of vertices is odd in this case). Let G^S and $G^{\bar{S}}$ be the graphs obtained from G by shrinking S and $\bar{S} = V \setminus S$ to a single node (see Figure 4.6). Let x^S and $x^{\bar{S}}$ be the restriction of x' to the edges of G^S and $G^{\bar{S}}$, respectively. By construction, $x^i(\delta(S)) = x^i(\delta(\bar{S}) = 1$ for $i = S, \bar{S}$.

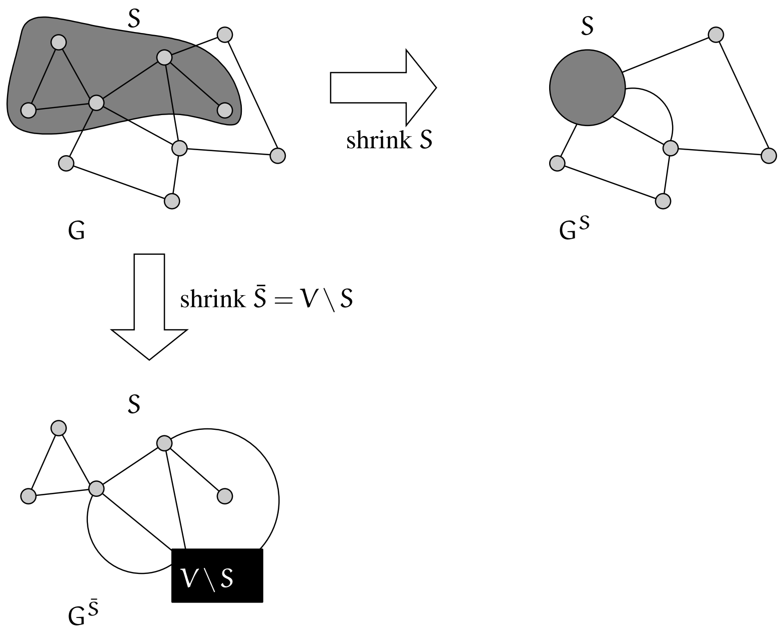


Figure 4.6: Graphs G^S and $G^{\bar{S}}$ obtained from G by shrinking the odd set S and $V \setminus S$ in the proof of Theorem 4.23.

It is easy to see that x^S and $x^{\bar{S}}$ satisfy the constraints (4.17) with respect to G^S and $G^{\bar{S}}$, respectively. Thus, by the induction hypothesis, we have $x^i \in \text{conv}(\text{PM}(G^i))$ for $i = S, \bar{S}$.

Hence, we can write x^i as convex combinations of perfect matchings of G^i:

$$x^S = \frac{1}{k}\sum_{j=1}^{k} \chi^{M_j^S} \tag{4.19}$$

$$x^{\bar{S}} = \frac{1}{k}\sum_{j=1}^{k} \chi^{M_j^{\bar{S}}} \tag{4.20}$$

Here, we have assumed without loss of generality that in both convex combinations the number of vectors used is the same, namely k. Also, we have assumed a special form of the convex combination which can be justified as follows: x' is an extreme point of P and thus is rational. This implies that x^i, $i = S, \bar{S}$ are also rational. Since all λ_j are rational, any convex combination $\sum_j \lambda_j y^j$ can be written by using common denominator k as $\sum_j \frac{\mu_j}{k} y^j$, where all μ_j are integral. Repeating vector y^j exactly μ_j times, we get the form $\frac{1}{k}\sum_j z^j$.

For $e \in \delta(S)$ the number of j such that $e \in M_j^S$ is $kx_e^S = kx'_e = kx_e^{\bar{S}}$. This is the same number of j such that $e \in M_j^{\bar{S}}$. Again: for every $e \in \delta(S)$ the number of j such that $e \in M_j^S$ is the same as the number of j with $e \in M_j^{\bar{S}}$. Thus, we can order the M_j^i so that M_j^S and $M_j^{\bar{S}}$ share an edge in $\delta(S)$ (any M_j^i, $i = S, \bar{S}$ has exactly one edge from $\delta(S)$). Then, $M_j := M_j^S \cup M_j^{\bar{S}}$ is a perfect matching of G since every vertex in G is matched and no vertex has more than one edge incident with it.

Let $M_j := M_j^S \cup M_j^{\bar{S}}$. Then we have:

$$x' = \frac{1}{k}\sum_{i=1}^{k} \chi^{M_j}. \tag{4.21}$$

Since M_j is a perfect matching of G we see from (4.21) that x' is a convex combination of perfect matchings of G. Since x' is an extreme point, it follows that x' must be a perfect matching itself. □

4.7 Total Dual Integrality

Another concept for proving integrality of a polyhedron is that of *total dual integrality*.

Definition 4.24 (Totally dual integral system)
*A rational linear system $Ax \leqslant b$ is **totally dual integral (TDI)**, if for each integral vector c such that*

$$z^{LP} = \max\left\{c^T x : Ax \leqslant b\right\}$$

is finite, the dual

$$\min\left\{b^T y : A^T y = c, y \geqslant 0\right\}$$

has an integral optimal solution.

Theorem 4.25 *If $Ax \leqslant b$ is TDI and b is integral, then the polyhedron $P = \{x : Ax \leqslant b\}$ is integral.*

Proof: If $Ax \leqslant b$ is TDI and b is integral, then the dual Linear Program

$$\min\left\{b^T y : A^T y = c, y \geqslant 0\right\}$$

has an optimal integral objective value if it is finite (the optimal vector y^* is integral and b is integral by assumption, so $b^T y^*$ is also integral). By LP-duality, we see that the value

$$z^{LP} = \max\left\{c^T x : Ax \leqslant b\right\}$$

is integral for all $c \in \mathbb{Z}^n$ where the value is finite. By Theorem 4.3(iv), the polyhedron P is integral. □

Example 4.26
Let $G = (A \cup B, E)$ be a complete bipartite graph. Suppose we wish to solve a generalized form of STABLESET on G in which we are allowed to pick a vertex more than once. Given weights c_{ab} for the edges (a,b) we want to solve the following Linear Program:

$$\max \sum_{v \in V} w_v x_v \tag{4.22a}$$

$$x_a + x_b \leqslant c_{ab} \qquad \text{for all } (a,b) \in E \tag{4.22b}$$

We claim that the system of ineqalities $x_a + x_b \leqslant c_{ab}$ $(a,b) \in A$ is TDI. The dual of (4.22) is given by:

$$\min \sum_{(a,b) \in E} c_{ab} y_{ab} \tag{4.23a}$$

$$\sum_{b \in B} y_{ab} = w_a \qquad \text{for all } a \in A \tag{4.23b}$$

$$\sum_{a \in A} y_{ab} = w_b \qquad \text{for all } b \in B \tag{4.23c}$$

$$y_{ab} \geqslant 0 \qquad \text{for all } (a,b) \in A. \tag{4.23d}$$

The constraint matrix of (4.23) is the constraint matrix of the assignment problem, which we have already shown to be totally unimodular (see Example 4.15). Thus, if the weight vector w is integral, then the dual (4.23) has an optimal integral solution (if it is feasible). ◁

It should be noted that the condition "and b is integral" in the previous theorem is crucial. It can be shown that for any rational system $Ax \leqslant b$ there is an integer ω such that $(1/\omega)Ax \leqslant (1/\omega)b$ is TDI. Hence, the fact that a system is TDI does not yet tell us anything useful about the structure of the corresponding polyhedron.

We have seen that if we find a TDI system with integral right hand side, then the corresponding polyhedron is integral. The next theorem shows that the converse is also true: if our polyhedron is integral, then we can also find a TDI system with integral right hand side defining it.

Theorem 4.27 *Let P be a rational polyhedron. Then, there exists a TDI system $Ax \leqslant b$ with A integral such that $P = P(A,b)$. Moreover, if P is an integral polyhedron, then b can be chosen to be integral.*

Proof: Let $P = \{x \in \mathbb{R}^n : Mx \leqslant d\}$ be a rational polyhedron. If $P = \emptyset$, then the claim is trivial. Thus, we assume from now on that P is nonempty. Since we can scale the rows of M by multiplying with arbitrary scalars, we can assume without loss of generality that M has integer entries. We also assume that the system $Mx \leqslant d$ does not contain any redundant rows. Thus, for any row m_i^T of M there is an $x \in P$ with $m_i^T x = d_i$.

Let $S = \left\{s \in \mathbb{Z}^n : s = M^T y, 0 \leqslant y \leqslant 1\right\}$ be the set of integral vectors which can be written as nonnegative linear combinations of the rows of M where no coefficient is larger than one.

Since y comes from a bounded domain and S contains only integral points, it follows that S is finite. For $s \in S$ we define

$$z(s) := \max\left\{s^T x : x \in P\right\}.$$

Observe that, if $s \in S$, say $s = \sum_{i=1}^{m} y_i m_i$, and $x \in P$, then $m_i^T x \leqslant d_i$ which means $y_i m_i^T x \leqslant y_i d_i$ for $i = 1, \dots, m$, from which we get that

$$s^T x = \sum_{i=1}^{m} y_i m_i^T x \leqslant \sum_{i=1}^{m} y_i d_i \leqslant \sum_{i=1}^{m} |d_i|.$$

Thus, $s^T x$ is bounded on P and $z(s) < +\infty$ for all $s \in S$. Moreover, the inequality $s^T x \leqslant z(s)$ is valid for P. We define the system $Ax \leqslant b$ to consist of all inequalities $s^T x \leqslant z(s)$ with $s \in S$.

Every row m_i^T of M is a vector in S (by assumption M is integral and m_i^T is a degenerated linear combination of the rows, namely with coefficient one for itself and zero for all other rows). Since $m_i^T x \leqslant d_i$ for all $x \in P$, the inequality $m_i^T x \leqslant d_i$ is contained in $Ax \leqslant b$. Furthermore, since we have only added valid inequalities to the system, it follows that

$$P = \{x : Ax \leqslant b\}. \tag{4.24}$$

If P is integral, then by Theorem 4.3 the value $z(s)$ is integral for each $s \in S$, so the system $Ax \leqslant b$ has an integral right hand side. The only thing that remains to show is that $Ax \leqslant b$ is TDI.

Let c be an integral vector such that $z^{LP} = \max\left\{c^T x : Ax \leqslant b\right\}$ is finite. We have to construct an optimal integral solution to the dual

$$\min\left\{b^T y : A^T y = c, y \geqslant 0\right\}. \tag{4.25}$$

We have

$$\begin{aligned}
z^{LP} &= \max\left\{c^T x : Ax \leqslant b\right\} \\
&= \max\left\{c^T x : x \in P\right\} && \text{(by (4.24))} \\
&= \max\left\{c^T x : Mx \leqslant d\right\} \\
&= \min\left\{d^T y : M^T y = c, y \geqslant 0\right\} && \text{(by LP-duality).} \qquad (4.26)
\end{aligned}$$

Let y^* be an optimal solution for the problem in (4.26) and consider the vector $\bar{s} = M^T(y^* - \lfloor y^* \rfloor)$. Observe that $y - \lfloor y^* \rfloor$ has entries in $[0, 1]$. Morever $\bar{s}$ is integral, since $\bar{s} = M^T y^* - M^T \lfloor y^* \rfloor = c - M^T \lfloor y^* \rfloor$ and c, M^T and $\lfloor y^* \rfloor$ are all integral. Thus, $\bar{s} \in S$. Now,

$$\begin{aligned}
z(\bar{s}) &= \max\left\{\bar{s}^T x : x \in P\right\} \\
&= \min\left\{d^T y : M^T y = \bar{s}, y \geqslant 0\right\} && \text{(by LP-duality).} \qquad (4.27)
\end{aligned}$$

The vector $y^* - \lfloor y^* \rfloor$ is feasible for (4.27) by construction. If v is feasible for (4.27), then $v + \lfloor y^* \rfloor$ is feasible for (4.26). Thus, it follows easily that $y - \lfloor y^* \rfloor$ is optimal for (4.27). Thus, $z(\bar{s}) = d^T(y^* - \lfloor y^* \rfloor)$, or

$$z^{LP} = d^T y^* = z(\bar{s}) + d^T \lfloor y^* \rfloor. \tag{4.28}$$

Consider the integral vector $\bar{y}$ defined as $\lfloor y^* \rfloor$ for the dual variables corresponding to rows in M, one for the dual variable corresponding to the constraint $\bar{s}^T x \leqslant z(\bar{s})$ and zero everywhere else. Clearly, $\bar{y} \geqslant 0$. Moreover,

$$A^T \bar{y} = \sum_{s \in S} \bar{y}_s s = M^T \lfloor y^* \rfloor + 1 \cdot \bar{s} = M^T \lfloor y^* \rfloor + M^T (y^* - \lfloor y^* \rfloor) = M^T y^* = c.$$

Hence, $\bar{y}$ is feasible for (4.25). Furthermore,

$$b^T \bar{y} = z(\bar{s}) + d^T \lfloor y^* \rfloor \stackrel{(4.28)}{=} z^{LP}.$$

Thus, $\bar{y}$ is an optimal integral solution for the dual (4.25). □

4.8 Submodularity and Matroids

In this section we apply our results about TDI systems to prove integrality for a class of important polyhedra.

Definition 4.28 (Submodular function)
*Let N be a finite set. A function $f\colon 2^N \to \mathbb{R}$ is called **submodular**, if*

$$f(A) + f(B) \geqslant f(A \cap B) + f(A \cup B) \text{ for all } A, B \subseteq N. \tag{4.29}$$

*The function is called **nondecreasing** if*

$$f(A) \leqslant f(B) \text{ for all } A, B \subseteq N \text{ with } A \subseteq B. \tag{4.30}$$

Usually we will not be given f "explicitly", that is, by a listing of all the $2^{|N|}$ pairs $(A, f(A))$. Rather, we will have access to f via an "oracle", that is, given A we can compute $f(A)$ by a call to the oracle.

Example 4.29
(i) The function $f(A) = |A|$ is nondecreasing and submodular.

(ii) Let $G = (V, E)$ be an undirected graph with edge weights $u\colon E \to \mathbb{R}_+$. The function $f\colon 2^V \to \mathbb{R}_+$ defined by $f(A) := \sum_{e \in \delta(A)} u(e)$ is submodular but not necessarily nondecreasing.

◁

Definition 4.30 (Submodular polyhedron, submodular optimization problem)
*Let f be submodular and nondecreasing. The **submodular polyhedron** associated with f is*

$$P(f) := \left\{ x \in \mathbb{R}^n_+ : \sum_{j \in S} x_j \leqslant f(S) \text{ for all } S \subseteq N. \right\} \tag{4.31}$$

*The **submodular optimization problem** is to optimize a linear objective function over $P(f)$:*

$$\max \left\{ c^T x : x \in P(f) \right\}.$$

Observe that by the polynomial time equivalence of optimization and separation (see Section 5.4) we can solve the submodular optimization problem in polynomial time if we can solve the corresponding separation problem in polynomial time: We index the polyhedra by the finite sets N, and it is easy to see that this class is proper.

We now consider the simple Greedy algorithm for the submodular optimization problem described in Algorithm 4.1. The surprising result proved in the following theorem is that the Greedy algorithm in fact solves the submodular optimization problem. But the result is even stronger:

Algorithm 4.1 Greedy algorithm for the submodular optimization problem.

GREEDY-SUBMODULAR

Sort the variables such that $c_1 \geqslant c_2 \geqslant \cdots \geqslant c_k > 0 \geqslant c_{k+1} \geqslant \cdots \geqslant c_n$.
Set $x_i := f(S^i) - f(S^{i-1})$ for $i = 1, \ldots, k$ and $x_i = 0$ for $i = k+1, \ldots, n$, where $S^i = \{1, \ldots, i\}$ and $S^0 = \varnothing$.

Theorem 4.31 *Let* f *be a submodular and nondecreasing function with* $f(\varnothing) = 0$, $c\colon N \to \mathbb{R}$ *be an arbitrary weight vector.*

(i) The Greedy algorithm solves the submodular optimization problem for maximizing $c^T x$ *over* $P(f)$.

(ii) The system (4.31) is TDI.

(iii) For integral valued f, *the polyhedron* $P(f)$ *is integral.*

Proof:

(i) Since f is nondecreasing, we have $x_i = f(S^i) - f(S^{i-1}) \geqslant 0$ for $i = 1, \ldots, k$. Let $S \subseteq N$. We have to show that $\sum_{j \in S} x_j \leqslant f(S)$. By the submodularity of f we have for $j \in S$:

$$\begin{aligned}
&f(\overbrace{S^j \cap S}^{=A}) + f(\overbrace{S^{j-1}}^{=B}) \geqslant f(\overbrace{\ S^j\ }^{=A\cup B}) + f(\overbrace{S^{j-1} \cap S}^{=A\cap B}) \\
\Leftrightarrow\ & f(S^j) - f(S^{j-1})\hat{A}\ \leqslant f(S^j \cap S) - f(S^{j-1} \cap S)
\end{aligned} \tag{4.32}$$

Thus,

$$\begin{aligned}
\sum_{j \in S} x_j &= \sum_{j \in S \cap S^k} \Big(f(S^j) - f(S^{j-1}) \Big) \\
&\leqslant \sum_{j \in S \cap S^k} \Big(f(S^j \cap S) - f(S^{j-1} \cap S) \Big) && \text{(by (4.32))} \\
&\leqslant \sum_{j \in S^k} \Big(f(S^j \cap S) - f(S^{j-1} \cap S) \Big) && \text{(f nondecreasing)} \\
&= f(S^k \cap S) - f(\varnothing) \\
&\leqslant f(S).
\end{aligned}$$

Thus, the vector x computed by the Greedy algorithm is in fact contained in $P(f)$. Its solution value is

$$c^T x = \sum_{i=1}^{k} c_i \Big(f(S^i) - f(S^{i-1}) \Big). \tag{4.33}$$

We now consider the Linear Programming dual of the submodular optimization problem:

$$\begin{aligned}
w^D = \min \quad & \sum_{S \subseteq N} f(S) y_S \\
& \sum_{S : j \in S} y_S \geqslant c_j && \text{for all } j \in N \\
& y_S \geqslant 0 && \text{for all } S \subseteq N.
\end{aligned}$$

If we can show that $c^T x = w^D$, then it follows that x is optimal. Construct a vector y by $y_{S^i} = c_i - c_{i+1}$ for $i = 1, \dots, k-1$, $y_{S^k} = c_k$ and $y_S = 0$ for all other sets $S \subseteq N$. Since we have sorted the sets such that $c_1 \geqslant c_2 \geqslant c_k > 0 \geqslant c_{k+1} \geqslant \cdots \geqslant c_n$, it follows that y has only nonnegative entries.

For $j = 1, \dots, k$ we have

$$\sum_{S: j \in S} y_S \geqslant \sum_{i=j}^{k} y_{S^i} = \sum_{i=j}^{k-1} (c_i - c_{i+1}) + c_k = c_j.$$

On the other hand, for $j = k+1, \dots, n$

$$\sum_{S: j \in S} y_S \geqslant 0 \geqslant c_j.$$

Hence, y is feasible for the dual. The objective function value for y is:

$$\begin{aligned}\sum_{S \subseteq N} f(S) y_S = \sum_{i=1}^{k} f(S^i) y_{S^i} &= \sum_{i=1}^{k-1} f(S^i)(c_i - c_{i+1}) + f(S^k) c_k \\ &= \sum_{i=1}^{k} (f(S^i) - f(S^{i-1}) c_i \\ &= c^T x,\end{aligned}$$

where the last equality stems from (4.33). Thus, y must be optimal for the dual and x optimal for the primal.

(ii) The proof of statement (ii) follows from the observation that, if c is integral, then the optimal vector y constructed is integral.

(iii) Follows from (ii) and Theorem 4.25.

□

An important class of submodular optimization problems are induced by special submodular functions, namely the rank functions of matroids.

Definition 4.32 (Independence system, matroid)
*Let N be a finite set and $\mathcal{I} \subseteq 2^N$. The pair $(N, \mathcal{I})$ is called an **independence system**, if $A \in \mathcal{I}$ and $B \subseteq A$ implies that $B \in \mathcal{I}$. The sets in $\mathcal{I}$ are called **independent sets**.*

*The independence system is a **matroid** if for each $A \in \mathcal{I}$ and $B \in \mathcal{I}$ with $|B| > |A|$ there exists $a \in B \setminus A$ with $A \cup \{a\} \in \mathcal{I}$.*

*Given a matroid $(N, \mathcal{I})$, its **rank function** $r \colon 2^N \to \mathbb{N}$ is defined by*

$$r(A) := \max\{|I| : I \subseteq A \textit{ and } I \in \mathcal{I}\}.$$

Observe that $r(A) \leqslant |A|$ for any $A \subset N$ and $r(A) = |A|$ if any only if $A \in \mathcal{I}$. Thus, we could alternatively specify a matroid $(N, \mathcal{I})$ also by (N, r).

Lemma 4.33 *The rank function of a matroid is submodular and nondecreasing.*

Proof: The fact that r is nondecreasing is trivial. Let $A, B \subseteq N$. We must prove that

$$r(A) + r(B) \geqslant r(A \cup B) + r(A \cap B).$$

Let $X \subseteq A \cup B$ with $|X| = r(A \cup B)$ and $Y \subseteq A \cap B$ with $|Y| = r(A \cap B)$. Let $X' := Y$. Since X' is independent and X is independent, if $|X'| < |X|$ we can add an element from X to X' without loosing independence. Continuing this procedure, we find X' with $|X'| = |X|$ and $Y \subseteq X'$ by construction. Hence, we can assume that $Y \subseteq X$. Now,

$$\begin{aligned} r(A) + r(B) &\geqslant |X \cap A| + |X \cap B| \\ &= |X \cap (A \cap B)| + |X \cap (A \cup B)| \\ &\geqslant |Y| + |X| \\ &= r(A \cap B) + r(A \cup B). \end{aligned}$$

This shows the claim. □

Example 4.34 (Matrix matroid)
Let A be an $m \times n$-matrix with columns $a_1, \dots, a_n$. Set $N := \{1, \dots, n\}$ and the family $\mathcal{I}$ by the condition that $S \in \mathcal{I}$ if and only if the vectors $\{a_i : i \in S\}$ are linearly independent. Then $(N, \mathcal{I})$ is an independendence system. By Steinitz' Theorem (basis exchange) from Linear algebra, we know that $(N, \mathcal{I})$ is in fact also a matroid. ◁

Example 4.35
Let $E = \{1, 3, 5, 9, 11\}$ and $\mathcal{F} := \left\{ A \subseteq E : \sum_{e \in A} e \leqslant 20 \right\}$. Then, $(E, \mathcal{F})$ is an independence system but not a matroid.

The fact that $(E, \mathcal{F})$ is an independence system follows from the property that, if $B \subseteq A \in \mathcal{F}$, then $\sum_{e \in B} e \leqslant \sum_{e \in A} e \leqslant 20$.

Now consider $B := \{9, 11\} \in \mathcal{F}$ and $A := \{1, 3, 5, 9\} \in \mathcal{F}$ where $|B| < |A|$. However, there is no element in $A \setminus B$ that can be added to B without losing independence. ◁

Definition 4.36 (Tree, forest)
*Let $G = (V, E)$ a graph. A **forest** in G is a subgraph (V, E') which does not contain a cycle. A **tree** is a forest which is connected (i.e., which contains a path between any two vertices).*

Example 4.37 (Graphic matroid)
Let $G = (V, E)$ be a graph. We consider the pair $(E, \mathcal{F})$, where

$$\mathcal{F} := \{ T \subseteq E : \ (V, T) \text{ is a forest} \}.$$

Clearly, $(E, \mathcal{F})$ is an independence system, since by deleting edges from a forest, we obtain again a forest. We show that the system is also a matroid. If (V, T) is a forest, then it follow easily by induction on $|T|$ that (V, T) has exactly $|V| - |T|$ connected components. Let $A \in \mathcal{I}$ and $B \in \mathcal{I}$ with $|B| > |A|$. Let $C_1, \dots, C_k$ be the connected components of (V, A) where $k = |V| - |A|$. Since (V, B) has fewer connected components, there must be an edge $e \in B \setminus A$ whose endpoints are in different components of (V, A). Thus $A \cup \{e\}$ is also a forest. ◁

Lemma 4.38 *Let $G = (V, E)$ be a graph and $(N, \mathcal{I})$ be the associated graphic matroid. Then the rank function r is given by $r(A) = |V| - \text{comp}(V, A)$, where $\text{comp}(V, A)$ denotes the number of connected components of the graph (V, A).*

Proof: Let $A \subseteq E$. It is easy to see that in a matroid all maximal independent subsets of A have the same cardinality. Let $C_1, \dots, C_k$ be the connected components of (V, A). For $i = 1, \dots, k$ we can find a spanning tree of C_i. The union of these spanning trees is a forest with $\sum_{i=1}^{k} (|C_i| - 1) = |V| - k$ edges, all of which are in A. Thus, $r(A) \geqslant |V| - k$.

Assume now that $F \subseteq A$ is an independent set. Then, F can contain only edges of the components $C_1, \dots, C_k$. Since F is a forest, the restriction of F to any C_i is also a forest, which implies that $|F \cap C_i| \leqslant |C_i| - 1$. Thus, we have $|F| \leqslant |V| - k$ and hence $r(A) \leqslant |V| - k$. □

Definition 4.39 (Matroid optimization problem)
Given a matroid $(N, \mathcal{I})$ *and a weight function* $c\colon N \to \mathbb{R}$, *the* ***matroid optimization problem*** *is to find a set* $A \in \mathcal{I}$ *maximizing* $c(A) = \sum_{a \in A} c(a)$.

By Lemma 4.33 the polyhedron

$$P(N, \mathcal{I}) := \left\{ x \in \mathbb{R}^n_+ : \sum_{j \in S} x_j \leqslant r(S) \text{ for all } S \subseteq N \right\}$$

is a submodular polyhedron. Moreover, by the integrality of the rank function and Theorem 4.31(iii) $P(N, \mathcal{I})$ is integral. Finally, since $r(\{j\}) \leqslant 1$ for all $j \in N$, it follows that $0 \leqslant x \leqslant 1$ for all $x \in P(N, \mathcal{I})$. Thus, in fact we have:

$$P(N, \mathcal{I}) = \operatorname{conv}\left(\left\{ x \in \mathbb{B}^n : \sum_{j \in S} x_j \leqslant r(S) \text{ for all } S \subseteq N. \right\}\right) \tag{4.34}$$

With (4.34) it is easy to see that the matroid optimization problem reduces to the submodular optimization problem. By Theorem 4.31(i) the Greedy algorithm finds an optimal (integral) solution and thus solves the matroid optimization problem.

It is worthwhile to have a closer look at the Greedy algorithm in the special case of a submodular polyhedron induced by a matroid. Let again be $S^i = \{1, \dots, i\}$, $S^0 = \emptyset$. Since $r(S^i) - r(S^{i-1}) \in \{0, 1\}$ and the algorithm works as follows:

(i) Sort the variables such that $c_1 \geqslant c_2 \geqslant \cdots \geqslant c_k > 0 \geqslant c_{k+1} \geqslant \cdots \geqslant c_n$.

(ii) Start with $S = \emptyset$.

(iii) For $i = 1, \dots, k$, if $S \cup \{i\} \in \mathcal{I}$, then set $S := S \cup \{i\}$.