# 7

# Fundamental concepts and results on polyhedra, linear inequalities, and linear programming

In this chapter we first state a fundamental theorem on linear inequalities (Section 7.1), and next we derive as consequences some other important results, like the Finite basis theorem for cones and polytopes, the Decomposition theorem for polyhedra (Section 7.2), Farkas' lemma (Section 7.3), the Duality theorem of linear programming (Section 7.4), an affine form of Farkas' lemma (Section 7.6), Carathéodory's theorem (Section 7.7), and results for strict inequalities (Section 7.8). In Section 7.5 we give a geometrical interpretation of LP-duality. In Section 7.9 we study the phenomenon of *complementary slackness*.

***Each of the results in this chapter holds both in real spaces and in rational spaces.*** In the latter case, all numbers occurring (like matrix and vector entries, variables) are restricted to the rationals.

## 7.1. THE FUNDAMENTAL THEOREM OF LINEAR INEQUALITIES

The fundamental theorem is due to Farkas [1894, 1898a] and Minkowski [1896], with sharpenings by Carathéodory [1911] and Weyl [1935]. Its geometric content is easily understood in three dimensions.

**Theorem 7.1** (Fundamental theorem of linear inequalities). *Let $a_1, \ldots, a_m, b$ be vectors in n-dimensional space. Then:*

*either* I. *$b$ is a nonnegative linear combination of linearly independent vectors from $a_1, \ldots, a_m$;*

*or* II. *there exists a hyperplane* $\{x \mid cx = 0\}$, *containing* $t-1$ *linearly indep ent vectors from* $a_1, \ldots, a_m$, *such that* $cb < 0$ *and* $ca_1, \ldots, ca_m \geqslant 0$, w $t := rank\{a_1, \ldots, a_m, b\}$.

*Proof.* We may assume that $a_1, \ldots, a_m$ span the $n$-dimensional space.

Clearly, I and II exclude each other, as otherwise, if $b = \lambda_1 a_1 + \cdots +$ with $\lambda_1, \ldots, \lambda_m \geqslant 0$, we would have the contradiction

(1) $$0 > cb = \lambda_1 \cdot ca_1 + \cdots + \lambda_m \cdot ca_m \geqslant 0.$$

To see that at least one of I and II holds, choose linearly independent $a_{i_1}, \ldots$ from $a_1, \ldots, a_m$, and set $D := \{a_{i_1}, \ldots, a_{i_n}\}$. Next apply the following iteratio

(2)
- (i) Write $b = \lambda_{i_1} a_{i_1} + \cdots + \lambda_{i_n} a_{i_n}$. If $\lambda_{i_1}, \ldots, \lambda_{i_n} \geqslant 0$, we are in case I
- (ii) Otherwise, choose the smallest $h$ among $i_1, \ldots, i_n$ with $\lambda_h < 0$. $\{x \mid cx = 0\}$ be the hyperplane spanned by $D \setminus \{a_h\}$. We normal so that $ca_h = 1$. [Hence $cb = \lambda_h < 0$.]
- (iii) If $ca_1, \ldots, ca_m \geqslant 0$ we are in case II.
- (iv) Otherwise, choose the smallest $s$ such that $ca_s < 0$. Then repla by $(D \setminus \{a_h\}) \cup \{a_s\}$, and start the iteration anew.

We are finished if we have shown that this process terminates. Let $D_k$ de the set $D$ as it is in the $k$th iteration. If the process does not terminate, $D_k = D_l$ for some $k < l$ (as there are only finitely many choices for $D$). Let the highest index for which $a_r$ has been removed from $D$ at the end of or the iterations $k, k+1, \ldots, l-1$, say in iteration $p$. As $D_k = D_l$, we know th also has been added to $D$ in some iteration $q$ with $k \leqslant q < l$. So

(3) $$D_p \cap \{a_{r+1}, \ldots, a_m\} = D_q \cap \{a_{r+1}, \ldots, a_m\}.$$

Let $D_p = \{a_{i_1}, \ldots, a_{i_n}\}$, $b = \lambda_{i_1} a_{i_1} + \cdots + \lambda_{i_n} a_{i_n}$, and let $c'$ be the vector $c$ fo in (ii) of iteration $q$. Then we have the contradiction:

(4) $$0 > c'b = c'(\lambda_{i_1} a_{i_1} + \cdots + \lambda_{i_n} a_{i_n}) = \lambda_{i_1} c' a_{i_1} + \cdots + \lambda_{i_n} c' a_{i_n} > 0.$$

The first inequality was noted in (2) (ii) above. The last inequality follows f

(5) if $i_j < r$ then $\lambda_{i_j} \geqslant 0, c'a_{i_j} \geqslant 0$; if $i_j = r$ then $\lambda_{i_j} < 0, c'a_{i_j} < 0$ (as by (2) (ii), $r$ is the smallest index wi $< 0$; similarly, by (2) (iv) $r$ is the sm index with $c'a_r < 0$)

if $i_j > r$ then $c'a_{i_j} = 0$ (by (3) and (2) (ii)).

The above proof of this fundamental theorem also gives a fundam algorithm: it is a disguised form of the famous *simplex method*, with *Bl rule* incorporated—see Chapter 11 (see Debreu [1964] for a similar proof with a lexicographic rule).

## 7.2. CONES, POLYHEDRA, AND POLYTOPES

We first derive some geometrical consequences from the Fundamental theorem. A nonempty set $C$ of points in Euclidean space is called a *(convex) cone* if $\lambda x + \mu y \in C$ whenever $x, y \in C$ and $\lambda, \mu \geqslant 0$. A cone $C$ is *polyhedral* if

(6) $$C = \{x \mid Ax \leqslant 0\}$$

for some matrix $A$, i.e. if $C$ is the intersection of finitely many linear half-spaces. Here a *linear half-space* is a set of the form $\{x \mid ax \leqslant 0\}$ for some nonzero row vector $a$. The cone *generated by* the vectors $x_1, \dots, x_m$ is the set

(7) $$\operatorname{cone}\{x_1, \dots, x_m\} := \{\lambda_1 x_1 + \dots + \lambda_m x_m \mid \lambda_1, \dots, \lambda_m \geqslant 0\},$$

i.e. it is the smallest convex cone containing $x_1, \dots, x_m$. A cone arising in this way is called *finitely generated*.

It now follows from the Fundamental theorem that for cones the concepts of 'polyhedral' and 'finitely generated' are equivalent (Farkas [1898a, 1902], Minkowski [1896], Weyl [1935]).

**Corollary 7.1a** (Farkas–Minkowski–Weyl theorem). *A convex cone is polyhedral if and only if it is finitely generated.*

*Proof.* I. To prove sufficiency, let $x_1, \dots, x_m$ be vectors in $\mathbb{R}^n$. We show that $\operatorname{cone}\{x_1, \dots, x_m\}$ is polyhedral. We may assume that $x_1, \dots, x_m$ span $\mathbb{R}^n$ (as we can extend any linear half-space $H$ of lin.hull$\{x_1, \dots, x_m\}$ to a linear half-space $H'$ of $\mathbb{R}^n$ such that $H' \cap \text{lin.hull}\{x_1, \dots, x_m\} = H$). Now consider all linear half-spaces $H = \{x \mid cx \leqslant 0\}$ of $\mathbb{R}^n$ such that $x_1, \dots, x_m$ belong to $H$ and such that $\{x \mid cx = 0\}$ is spanned by $n-1$ linearly independent vectors from $x_1, \dots, x_m$. By Theorem 7.1, $\operatorname{cone}\{x_1, \dots, x_m\}$ is the intersection of these half-spaces. Since there are only finitely many such half-spaces, the cone is polyhedral.

II. Part I above also yields the converse implication. Let $C$ be a polyhedral cone, say $C = \{x \mid a_1^{\mathrm{T}} x \leqslant 0, \dots, a_m^{\mathrm{T}} x \leqslant 0\}$ for certain column vectors $a_1, \dots, a_m$. As by I above, each finitely generated cone is polyhedral, there exist column vectors $b_1, \dots, b_t$ such that

(8) $$\operatorname{cone}\{a_1, \dots, a_m\} = \{x \mid b_1^{\mathrm{T}} x \leqslant 0, \dots, b_t^{\mathrm{T}} x \leqslant 0\}.$$

We show that $C = \operatorname{cone}\{b_1, \dots, b_t\}$, implying that $C$ is finitely generated.

Indeed, $\operatorname{cone}\{b_1, \dots, b_t\} \subseteq C$, as $b_1, \dots, b_t \in C$, since $b_j^{\mathrm{T}} a_i \leqslant 0$ for $i = 1, \dots, m$ and $j = 1, \dots, t$, by (8). Suppose $y \notin \operatorname{cone}\{b_1, \dots, b_t\}$ for some $y \in C$. By Part I, $\operatorname{cone}\{b_1, \dots, b_t\}$ is polyhedral, and hence there exists a vector $w$ such $w^{\mathrm{T}} b_1, \dots, w^{\mathrm{T}} b_t \leqslant 0$ and $w^{\mathrm{T}} y > 0$. Hence by (8), $w \in \operatorname{cone}\{a_1, \dots, a_m\}$, and hence $w^{\mathrm{T}} x \leqslant 0$ for all $x$ in $C$. But this contradicts the facts that $y$ is in $C$ and $w^{\mathrm{T}} y > 0$. □

A set $P$ of vectors in $\mathbb{R}^n$ is called a *(convex) polyhedron* if

(9) $$P = \{x \mid Ax \leqslant b\}$$

for some matrix $A$ and vector $b$, i.e. if $P$ is the intersection of finitely many affine half-spaces. Here an *affine half-space* is a set of the form $\{x \mid wx \leqslant \delta\}$ for some nonzero row vector $w$ and some number $\delta$. If (9) holds, we say that $Ax \leqslant b$ *defines* or *determines* $P$. Trivially, each polyhedral cone is a polyhedron.

A set of vectors is a (*convex*) *polytope* if it is the convex hull of finitely many vectors.

It is intuitively obvious that the concepts of polyhedron and of polytope are related. This is made more precise in the Decomposition theorem for polyhedra (Motzkin [1936]) and in its direct corollary, the Finite basis theorem for polytopes (Minkowski [1896], Steinitz [1916], Weyl [1935]).

**Corollary 7.1b** (Decomposition theorem for polyhedra). *A set $P$ of vectors in Euclidean space is a polyhedron, if and only if $P = Q + C$ for some polytope $Q$ and some polyhedral cone $C$.*

*Proof.* I. First let $P = \{x \mid Ax \leqslant b\}$ be a polyhedron in $\mathbb{R}^n$. By Corollary 7.1a the polyhedral cone

(10) $$\left\{\begin{pmatrix} x \\ \lambda \end{pmatrix} \middle| x \in \mathbb{R}^n; \lambda \in \mathbb{R}; \lambda \geqslant 0; Ax - \lambda b \leqslant 0\right\}$$

is generated by finitely many vectors, say by $\begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_m \\ \lambda_m \end{pmatrix}$. We may assume that each $\lambda_i$ is 0 or 1. Let $Q$ be the convex hull of the $x_i$ with $\lambda_i = 1$, and let $C$ be the cone generated by the $x_i$ with $\lambda_i = 0$. Now $x \in P$, if and only if $\begin{pmatrix} x \\ 1 \end{pmatrix}$ belongs to (10), and hence, if and only if $\begin{pmatrix} x \\ 1 \end{pmatrix} \in \text{cone}\left\{\begin{pmatrix} x_1 \\ \lambda_1 \end{pmatrix}, \ldots, \begin{pmatrix} x_m \\ \lambda_m \end{pmatrix}\right\}$. It follows directly that $P = Q + C$.

II. Let $P = Q + C$ for some polytope $Q$ and some polyhedral cone $C$. Say $Q = \text{conv.hull}\,\{x_1, \ldots, x_m\}$ and $C = \text{cone}\,\{y_1, \ldots, y_t\}$. Then a vector $x_0$ belongs to $P$, if and only if

(11) $$\begin{pmatrix} x_0 \\ 1 \end{pmatrix} \in \text{cone}\left\{\begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} x_m \\ 1 \end{pmatrix}, \begin{pmatrix} y_1 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} y_t \\ 0 \end{pmatrix}\right\}.$$

By Corollary 7.1a, the cone in (11) is equal to $\left\{\begin{pmatrix} x \\ \lambda \end{pmatrix} \middle| Ax + \lambda b \leqslant 0\right\}$ for some matrix $A$ and vector $b$. Hence $x_0 \in P$, if and only if $Ax_0 \leqslant -b$, and therefore $P$ is a polyhedron. □

We shall say that $P$ is *generated by the points* $x_1, \ldots, x_m$ and *by the directions* $y_1, \ldots, y_t$ if.

(12) $$P = \text{conv.hull}\,\{x_1, \ldots, x_m\} + \text{cone}\,\{y_1, \ldots, y_t\}.$$

This gives a 'parametric' description of the solution set of a system of linear inequalities. For more about decomposition of polyhedra, see Section 8.9 below.

The Finite basis theorem for polytopes can be derived from the Decomposition theorem. It is usually attributed to Minkowski [1896], Steinitz [1916] and Weyl [1935]. ('This classical result is an outstanding example of a fact which is completely obvious to geometric intuition, but which wields important algebraic content and is not trivial to prove'—R.T. Rockafellar.)

**Corollary 7.1c** (Finite basis theorem for polytopes). *A set $P$ is a polytope if and only if $P$ is a bounded polyhedron.*

*Proof.* Directly from Corollary 7.1b. □

## 7.3. FARKAS' LEMMA AND VARIANTS

Another consequence of Theorem 7.1 is the well-known Farkas' lemma, proved first by Farkas [1894, 1898a] and Minkowski [1896].

**Corollary 7.1d** (Farkas' lemma). *Let $A$ be a matrix and let $b$ be a vector. Then there exists a vector $x \geqslant 0$ with $Ax = b$, if and only if $yb \geqslant 0$ for each row vector $y$ with $yA \geqslant 0$.*

*Proof.* The necessity of the condition is trivial, as $yb = yAx \geqslant 0$ for all $x$ and $y$ with $x \geqslant 0$, $yA \geqslant 0$, and $Ax = b$. To prove sufficiency, suppose there is no $x \geqslant 0$ with $Ax = b$. Let $a_1, \ldots, a_m$ be the columns of $A$. Then $b \notin \text{cone}\{a_1, \ldots, a_m\}$, and hence, by Theorem 7.1, $yb < 0$ for some $y$ with $yA \geqslant 0$. □

Farkas' lemma is equivalent to: if the linear inequalities $a_1x \leqslant 0, \ldots, a_nx \leqslant 0$ imply the linear inequality $wx \leqslant 0$, then $w$ is a nonnegative linear combination of $a_1, \ldots, a_n$ (thus providing a 'proof' of the implication).

Geometrically, Farkas' lemma is obvious: the content is that if a vector $b$ does not belong to the cone generated by the vectors $a_1, \ldots, a_n$, there exists a linear hyperplane separating $b$ from $a_1, \ldots, a_n$.

There are several other, equivalent, forms of Farkas' lemma, like those described by the following two corollaries.

**Corollary 7.1e** (Farkas' lemma (variant)). *Let $A$ be a matrix and let $b$ be a vector. Then the system $Ax \leqslant b$ of linear inequalities has a solution $x$, if and only if $yb \geqslant 0$ for each row vector $y \geqslant 0$ with $yA = 0$.*

*Proof.* Let $A'$ be the matrix $[I \;\; A \;\; -A]$. Then $Ax \leqslant b$ has a solution $x$, if and only if $A'x' = b$ has a nonnegative solution $x'$. Application of Corollary 7.1d to the latter system yields Corollary 7.1e. □

[Kuhn [1956a] showed that this variant of Farkas' lemma can be proved in a nice short way with the 'Fourier–Motzkin elimination method'—see Section 12.2.]

**NOPT048: Optimalizační metody - LS 2012** Jiří Sgall

Stěny mnohostěnu, minimální popis mnohostěnu

16. 3. 2012, 4. přednáška *Zapsal: Václav Obrázek*

Poslední změna: 16. června 2012

# 1 Mnohostěny

## 1.1 Stěna mnohostěnu

**Definice 1.** Nechť $P$ je konvexní mnohostěn a $c \in \mathbb{R}^n, t \in \mathbb{R}$. Jestliže $\forall x \in P : c^T x \leq t$ a $\exists x \in P : c^T x = t$, pak množinu $\{x | c^T x = t\}$ nazýváme **tečná nadrovina**, množiny $\{x | c^T x = t\} \cap P$, $\emptyset$ a $P$ nazýváme **stěnami** $P$. Stěnu $F$, pro kterou platí $F \neq P$ a $F \neq \emptyset$, nazveme **vlastní stěna**.

**Definice 2.** **Vrchol** $P$ je stěna dimenze 0.
**Hrana** $P$ je stěna dimenze 1.
**Faseta** $P$ je stěna dimenze $\dim(P) - 1$

**Věta 3.** *Průnik stěn $P$ je stěna $P$.*

*Důkaz.* Na cvičeních. □

**Věta 4.** *Nechť $P$ je konvexní mnohostěn, $V$ množina vrcholů $P$.*
*Pak* $x \in V \Leftrightarrow \underbrace{x \notin \text{conv}(P \setminus \{x\})}_{x \text{ je extremální bod}}$.
*Navíc, pokud je $P$ omezený, tak $P = \text{conv}(V)$.*

*Důkaz.* (Pro $V$ omezené)
$V_0 :=$ minimální množina (vzhledem k inkluzi), že $P = \text{conv}(V_0)$
$V_{ext} := \{x \in P | x \notin \text{conv}(P \setminus \{x\})\}$
Idea důkazu: $V \subseteq V_{ext} \subseteq V_0 \subseteq V \Rightarrow V_{ext} = V_0$

$V \subseteq V_{ext}$ Sporem, volíme $z \in V, z \notin V_{ext}$
Z definice vrcholu $z \in P, c^T z = t$ a $\forall x \in P \setminus \{z\} : c^T x < t$
$z \notin V_{ext} \Rightarrow z \in \text{conv}(\underbrace{\{x_1, \dots x_k\}}_{\subseteq P \setminus \{z\}}) \Rightarrow$ spor, $z$ není vrchol.

$V_{ext} \subseteq V_0$ Sporem, volíme $z \in V_{ext} \setminus V_0$
$z \in P = \text{conv}(V_0) \subseteq \text{conv}(P \setminus \{z\})$ spor

$V_0 \subseteq V$ Použijeme Větu o oddělování
$z \in V_0, D := \text{conv}(V_0 \setminus \{z\})$; $\{z\}$ a $D$ jsou disjunktní uzavřené konvexní množiny. Tedy podle věty o oddělování máme $c, r$ takové, že $\forall x \in D : c^T x < r$ a $c^T z > r$
Zvolíme $t := c^T z$, pak platí $\forall x \in D : c^T x < r < t$
a tedy $A = \{x | c^T x = t\}$ je tečná nadrovina, tedy $\forall x \in V_0 : c^T x \leq t$ a $\forall x \in P : c^T x \leq t$
Tvrdíme: $A \cap P = \{z\}$, nechť $z' \in A \cap P; z' \neq z$, pak

ectrical network □

# Chapter 4

# Matroids

Matroids provide a successful connection between graph theory, geometry and linear algebra. Some of the dualities we will discuss later are rooted in the theory of matroids. Moreover, matroids provide a basis for discrete optimization. Several important algorithms, for instance the greedy algorithm, belong to the matroid world. We make a notational agreement in this chapter: the graphs are allowed to have loops and multiple edges.

**Definition 4.0.5.** Let $X$ be a finite set and $S \subset 2^X$. We say that $M = (X, S)$ is a matroid if the following conditions are satisfied:

(I1) $\emptyset \in S$,

(I2) $A \in S$ and $A' \subset A$ then $A' \in S$ ($S$ is *hereditary*),

(I3) $U, V \in S$ and $|U| = |V| + 1$ then there is $x \in U - V$ so that $V \cup \{x\} \in S$ ($S$ satisfies an *exchange axiom*).

*Example* 4.0.6. Let $X$ be the set of all columns of a matrix over a field and let $S$ consist of all the subsets of $X$ that are linearly independent. Then $(X, S)$ is a matroid (called *vectorial or linear matroid*).

**Definition 4.0.7.** Let $M = (X, S)$ be a matroid. The elements of $S$ are called *independent sets* of $M$. The maximal elements of $S$ (w.r.t. inclusion) are called *bases.* Let $A \subset X$. The *rank* of $A$, $r(A)$, is defined by $r(A) = max\{|A'|; A' \subset A, A' \in S\}$. The *closure* of $A$, $\sigma(A)$, equals $\{x; r(A \cup \{x\}) = r(A)\}$. If $A = \sigma(A)$ then $A$ is *closed.*

By repeated use of (I3) in Definition 4.0.5 we get

**Corollary 4.0.8.** *If $U, V \in S$ and $|U| > |V|$ then there is $Z \subset U - V$, $|Z| = |U - V|$ and $V \cup Z \in S$. All bases have the same cardinality.*

**Theorem 4.0.9.** *A non-empty collection $\mathcal{B}$ of subsets of $X$ is the set of all bases of a matroid on $X$ if and only if the following condition is satisfied.*

(B1) *If* $B_1, B_2 \in \mathcal{B}$ *and* $x \in B_1 - B_2$ *then there is* $y \in B_2 - B_1$ *such that* $B_1 - \{x\} \cup \{y\} \in \mathcal{B}$.

*Proof.* Property (B1) is true for matroids: we apply (I3) to $B_1 - \{x\}, B_2$. To show the other implication we need to prove that each hereditary system satisfying (B1) satisfies (I3) too. First we observe that (B1) implies that no element of $\mathcal{B}$ is a strict subset of another one, and by repeated application of (B1) we observe that in fact all the elements of $\mathcal{B}$ have the same size. To show (I3) let $B_U, B_V$ be bases containing $U, V$ from (I3) and such that their symmetric difference is as small as possible. If $(B_V \cap (U - V)) \neq \emptyset$ then any element from there may be added to $V$ and (I3) holds. We show that $(B_V \cap (U - V)) = \emptyset$ leads to a contradiction with the choice of $B_U, B_V$: If $x \in B_V - B_U - V$ then (B1) produces a pair of bases with smaller symmetric difference. Hence $B_V - B_U - V$ is empty. But then necessarily $|B_V| < |B_U|$, a contradiction. □

**Theorem 4.0.10.** *A collection* $S$ *of subsets of* $X$ *is the set of all independent sets of a matroid on* $X$ *if and only if (I1), (I2) and the following condition are satisfied.*

(I3') *If* $A$ *is any subset of* $X$ *then all the maximal (w.r.t. inclusion) subsets* $Y$ *of* $A$ *with* $Y \in S$ *have the same cardinality.*

*Proof.* Property (I3') is clerly equivalent to (I3). □

**Theorem 4.0.11.** *An integer function* $r$ *on* $2^X$ *is a rank function of a matroid on* $X$ *if and only if the following conditions are satisfied.*

(R1) $r(\emptyset) = 0$,

(R2) $r(Y) \le r(Y \cup \{y\}) \le r(Y) + 1$,

(R3) *If* $r(Y \cup \{y\}) = r(Y \cup \{z\}) = r(Y)$ *then* $r(Y) = r(Y \cup \{y, z\})$.

*Proof.* Clearly (R1),(R2) hold for matroids. To show (R3) let $B$ be a maximal independent subset of $Y$. If $r(Y) < r(Y \cup \{y, z\})$ then $B$ is not maximal independent in $Y \cup \{y, z\}$, but any enlargement leads to a contradiction.
To show the other direction we say that $A$ is independent if $r(A) = |A|$. Obviously the set of the independent sets satisfies (I1). If $A$ is independent and $B \subset A$ then $r(B) = |B|$ since otherwise, by (R2), $r(A) \le |B - A| + r(B) < |A|$. Hence (I2) holds. If (I3) does not hold for $U, V$ then by repeated application of (R3) we get that $r(V \cup (U - V)) = r(V)$, but this set contains $U$, a contradiction. □

**Theorem 4.0.12.** *An integer function on* $2^X$ *is a rank function of a matroid on* $X$ *if and only if the following conditions are satisfied.*

(R1') $0 \le r(Y) \le |Y|$,

(R2') $Z \subset Y$ *implies* $r(Z) \le r(Y)$,

(R3') $r(Y \cup Z) + r(Y \cap Z) \le r(Y) + r(Z)$. *This property is called submodularity.*

*Proof.* Clearly (R1') and (R2') hold for matroids. To show (R3') let $B$ be a maximal independent set in $Y \cap Z$ and let $B_Y, B_Z$ be maximal independent in $Y, Z$ containing $B$. We have $r(Y \cap Z) = |B_Y \cap B_Z|$ and clearly $r(B_Y \cup B_Z) \leq |Y \cup Z|$. Hence (R3') follows. On the other hand, (R1),(R2) and (R3) follow easily from (R1'), (R2') and (R3'). □

**Theorem 4.0.13.** *The closure $\sigma(A)$ is the smallest (w.r.t. inclusion) closed set containing $A$.*

*Proof.* First observe that $\sigma(A)$ is closed, since $r(\sigma(A) \cup \{x\}) = r(\sigma(A))$ implies $r(A \cup \{x\}) \leq r(\sigma(A) \cup \{x\}) = r(\sigma(A)) = r(A)$. To show the second part let $A \subset C$, $C$ closed and $x \in (\sigma(A) - C)$. Hence $r(C \cup \{x\}) > r(C)$ and this implies $r(A \cup \{x\}) > r(A)$. (exercise: why?) This contradicts $x \in \sigma(A)$. □

**Theorem 4.0.14.** *A function $\sigma : 2^X \rightarrow 2^X$ is the closure operator of a matroid on $X$ if and only if the following conditions are satisfied.*

*(S1)* $Y \subset \sigma(Y)$,

*(S2)* $Z \subset Y$ *then* $\sigma(Z) \subset \sigma(Y)$,

*(S3)* $\sigma(\sigma(Y)) = \sigma(Y)$,

*(S4) if $y \notin \sigma(Y)$ but $y \in \sigma(Y \cup \{z\})$ then $z \in \sigma(Y \cup \{y\})$. This property is called the Steinitz-MacLane exchange axiom.*

We say that two matroids are isomorphic if they differ only in the names of their groundset elements.

## 4.1 Examples of matroids

We already know vectorial matroids. A matroid is *representable* if it is isomorphic to a vectorial matroid.
Let $G = (V, E)$ be a graph and let $M(G) = (E, S)$ where $S = \{F \subset E; F \text{ forest}\}$. Then $M(G)$ is a matroid, called the *cycle matroid* of $G$. Its rank function is $r(F) = |V| - c(F)$, where we recall that $c(F)$ denotes the number of connected components of the spanning subgraph $(V, F)$. The matroids isomorphic to cycle matroids of graphs are called *graphic matroids.*
Let $G = (V, E)$ be a graph. The *matching matroid* of $G$ is the pair $(V, S)$ where $A \in S$ if and only if $A$ may be covered by a matching of $G$. This is a matroid since the basis axiom corresponds to the exchange along an alternating path of two maximum matchings of $G$.

A matroid is *simple* if $r(A) = |A|$ whenever $|A| < 3$. Simple matroids of rank 3 have a natural representation that we now describe. Each matroid is determined by its rank function and so each simple matroid $M$ of rank 3 is determined by the set $L(M) = \{A \subset X; |A| > 2, r(A) = 2, A \text{ closed }\}$; if $|A| > 2$ then $r(A) = 2$ if and only if $A$ is a subset of an element of $L(M)$.

**Lemma 4.1.1.** *If $A, B \in L(M)$ then $|A \cap B| \leq 1$.*

*Proof.* We assume for a contradiction $\{x, z\} \subset A \cap B$, $a \in A - B$ and $b \in B - A$. Then both $a, b$ belong to $\sigma(\{x, z\})$ and hence, by Theorem 4.0.13, both $a, b$ belong to any closed set containing $\{x, z\}$: a contradiction.

□

A set $C \subset 2^X$ is a *configuration* on $X$ if each element of $C$ has at least 3 elements and any pair of elements of $C$ have at most one element of $X$ in common.

**Theorem 4.1.2.** *Each configuration is the set $L(M)$ of a simple matroid of rank 3 on $X$.*

*Proof.* Given $C$, for each $A \subset X$ define $r(A) = |A|$ if $|A| \leq 2$, and if $|A| > 2$ then $r(A) = 2$ if and only if $A$ is a subset of an element of $C$; $r(A) = 3$ otherwise. We show that $r$ is a rank function of a matroid. Note that $(R1, (R2)$ are obviously satisfied. We show $(R3)$: If $r(Y \cup \{y\}) = r(Y \cup \{z\}) = r(Y)$ then $|Y| \geq 2$ and both $Y \cup \{y\}$, $Y \cup \{z\}$ are subsets of an element of $C$. They are in fact subsets of the same element of $C$ since their intersection has size 2. Hence $r(Y) = r(Y \cup \{y, z\})$.

□

Hence we can represent simple matroids of rank 3 by a system of 'lines' in the plane corresponding to the elements of $L(M)$. The most famous picture of matroid theory, the *Fano matroid* $F_7$, is depicted in Figure 4.1. The Fano matroid is the vectorial matroid, over $GF(2)$, of the matrix whose columns are all non-zero vectors of $GF(2)^3$.

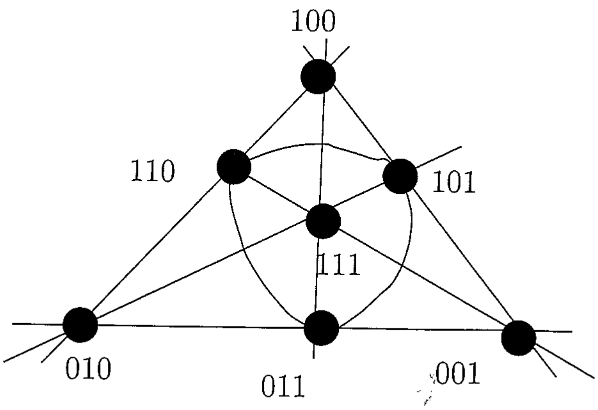


**Figure 4.1.** Fano matroid $F_7$

## 4.2 Greedy algorithm

Let $(X, S)$ be a set system and $w$ a weight function on $X = \{1, 2, \cdots, n\}$. In a *discrete optimization problem* we may want to find $J \in S$ such that $\sum_{i \in J} w_i$ is maximized. We encountered the *greedy algorithm* (GA) in Section 3.1. There, it was shown that GA correctly solves the minimum spanning tree problem (in order to turn the minimum spanning tree problem into a maximization problem we change the sign of each weight). Let us first define the greedy algorithm in a more general way, as an algorithm for the general optimisation problem that works as follows:

- Order the elements of $X$ so that $w_1 \geq w_2 \geq \cdots \geq w_n$.
- $J := \emptyset$.
- For $i = 1, \cdots, n$ do: if $J \cup \{i\} \in S$ and $w_i \geq 0$ then $J := J \cup \{i\}$.

The next theorem shows that applicability of GA characterizes matroids.

**Theorem 4.2.1.** *Let $(X, S)$ be a hereditary non-empty set system. Then the greedy algorithm solves the discrete optimization problem correctly for any weight function $w$ on $X$ if and only if $(X, S)$ is a matroid.*

*Proof.* If a hereditary system is not a matroid then it does not satisfy (I3') and it is not difficult to construct a weight function $w$ for which the greedy algorithm does not work. Let us prove the opposite implication: Let $m$ be maximal such that $w_m \geq 0$. Let $z'$ be the characteristic vector of a set produced by the greedy algorithm and let $z$ be the characteristic vector of any other set of $S$. Let $T_i = \{1, \cdots, i\}$, $i = 1, \cdots, m$. We notice that for each $i$

$$z'(T_i) = \sum_{j \leq i} z'_j \geq \sum_{j \leq i} z_j = z(T_i),$$

since $J \cap T_i$ is a maximal subset of $T_i$ which belongs to $S$ (by the definition of GA). We have

$$wz \leq \sum_{i=1}^{m} w_i z_i = \sum_{i=1}^{m} w_i (z(T_i) - z(T_{i-1})) =$$

$$\sum_{i=1}^{m-1} (w_i - w_{i+1}) z(T_i) + w_m z(T_m) \leq \sum_{i=1}^{m-1} (w_i - w_{i+1}) z'(T_i) + w_m z'(T_m) = wz'.$$

□

The only property we used in the proof is that $z \geq 0$ and $z(T_i) \leq z'(T_i) = r(T_i)$. GA thus solves also the following problem:

maximize $\sum_{i \in X} w_i z_i$

$z(A) = \sum_{i \in A} z_i \leq r(A)$, $A \subset X$;

$z_i \geq 0,\ i \in X.$

The problems that may be described in this form are called *linear programs*, and the part of optimization which studies linear programs is called *linear programming*.

**Corollary 4.2.2.** *Edmonds Matroid Polytope theorem: For any matroid, the convex hull of the characteristic vectors of the independent sets is equal to* $\mathcal{P} = \{z \geq 0;\ \text{for each } A \subset X, z(A) \leq r(A)\}$.

*Proof.* (sketch) The convex hull is clearly a subset of $\mathcal{P}$. By the Minkowski-Weyl theorem introduced in the beginning of the book we have that $\mathcal{P}$, a bounded intersection of finitely many half-spaces, is a *polytope*, i.e. a convex hull of its *vertices.* Each vertex $c$ of $\mathcal{P}$ is characterized by the existence of a half-space $\{z; wz \leq b\}$ which intersects $\mathcal{P}$ exactly in $\{c\}$. Since GA solves any problem $\max\{wz; z \in \mathcal{P}\}$, each non-empty intersection of $\mathcal{P}$ with a half-space necessarily contains the incidence vector of an independent set. In particular, each vertex of $\mathcal{P}$ is the incidence vector of an independent set, and the theorem follows. □

Finally we remark that the greedy algorithm is polynomial time if there is a polynomial algorithm to answer the questions 'Is $J$ independent ?'. It is usual for matroids to be given, for algorithmic purposes, by such an independence-testing oracle.

## 4.3 Circuits

**Definition 4.3.1.** A *circuit* in a matroid is a minimal (w.r.t. inclusion) non-empty dependent set.

The circuits of graphic matroids are the cycles of the underlying graphs.

**Theorem 4.3.2.** *A non-empty set $\mathcal{C}$ is the set of the circuits of a matroid if and only if the following conditions are satisfied.*

*(C1) If $C_1 \neq C_2$ are circuits then $C_1$ is not a subset of $C_2$,*

*(C2) If $C_1 \neq C_2$ are circuits and $z \in C_1 \cap C_2$ then $(C_1 \cup C_2) - z$ contains a circuit.*

*Proof.* First we show that a matroid satisfies the above properties. The first is obvious. For the second we have $r(C_1 \cup C_2) \leq r(C_1) + r(C_2) - r(C_1 \cap C_2) = |C_1| + |C_2| - |C_1 \cap C_2| - 2 = |C_1 \cup C_2| - 2$. Hence $(C_1 \cup C_2) - z$ must be dependent. On the other hand, we define $S$ to be the set of all subsets which do not contain an element of $\mathcal{C}$ and show that $(X, S)$ is a matroid. Axioms (I1) and (I2) are obvious and we show (I3'): let $A \subset X$ and for a contradiction let $J_1, J_2$ be maximal subsets of $A$ that belong to $S$ and $|J_1| < |J_2|$, and let $|J_1 \cap J_2|$ be as large as possible. Let $x \in J_1 - J_2$ and $C$ the unique circuit of $J_2 \cup x$. Necessarily there is $f \in C - J_1$ and $J_3 = (J_2 \cup x) - f$ belongs to $S$ by the uniqueness of $C$. Then $|J_3 \cap J_1| < |J_2 \cap J_1|$, a contradiction. □



**Corollary 4.3.3**

**Proposition 4.**
*is a circuit C wi*

*Proof.* If $x \in \sigma($
and hence conta
dent set in $A$ co
and hence $x \in \sigma$

## 4.4 Basic

**Definition 4.4.**
independent in

Each truncat

**Definition 4.4.**
sum of $M_1, M_2)$
only if $A \cap X_1$ is

**Definition 4.4.**
$\{A \subset X_i; |A| \leq$

It follows im
matroid. This o

**Definition 4.4**
$T' = X \setminus T$. $M$
independent if a

**Theorem 4.4.**
$r(A \cup T) - r(T)$

*Proof.* Obviousl
maximal subset
maximal indepe

## 4.5 Dual

**Definition 4.5**
$(X, S^*)$ such th

**Proposition 4**
$|A| - r(X) + r($

*Proof.* Again th
maximal subset
$M)$ subset of $X$
If there is $x \in$
$A \setminus J \subset B'$ and