Submodularity and Discrete Convexity

Satoru Fujishige

Research Institute for Mathematical Sciences
Kyoto University

4th Workshop on Discrete Optimization in Machine Learning (DISCML)

December 7, 2012
Lake Tahoe, Nevada, United States

E: a nonempty finite set

A submodular function $f : 2^E \to \mathbf{R}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \qquad (\forall X,\, Y \subseteq E)$$

E: a nonempty finite set

A submodular function $f : 2^E \to \mathbf{R}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \qquad (\forall X,\ Y \subseteq E)$$

Lemma: $\mathcal{D}_{\min}(f)$: the set of all minimizers of f

$$X,\ Y \in \mathcal{D}_{\min}(f) \implies X \cup Y,\ X \cap Y \in \mathcal{D}_{\min}(f)$$

i.e., $\mathcal{D}_{\min}(f)$ is a distributive lattice.

E: a nonempty finite set

A submodular function $f : 2^E \to \mathbf{R}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \qquad (\forall X,\ Y \subseteq E)$$

Lemma: $\mathcal{D}_{\min}(f)$: the set of all minimizers of f

$$X,\ Y \in \mathcal{D}_{\min}(f) \implies X \cup Y,\ X \cap Y \in \mathcal{D}_{\min}(f)$$

i.e., $\mathcal{D}_{\min}(f)$ is a distributive lattice.

It leads us to a Characterization of Submodular Functions

Theorem: A set function $f : 2^E \to \mathbf{R}$ is a submodular function

$\Updownarrow$

For $\forall$ modular function $\mu : 2^E \to \mathbf{R}$, $\mathcal{D}_{\min}(f - \mu)$ is a distributive lattice.

$\longrightarrow$

Distributive Lattices and Posets

Theorem (Birkhoff-Iri): $\mathcal{D} \subseteq 2^E$, $\emptyset, E \in \mathcal{D}$.

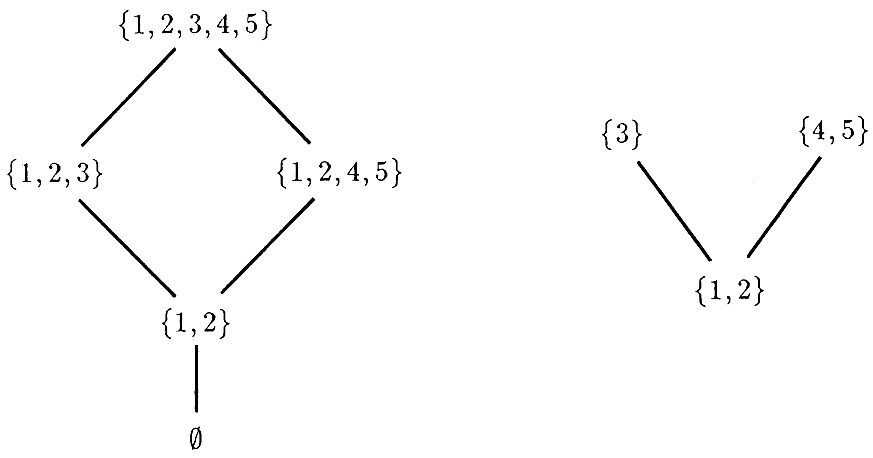
$\mathcal{D}$ is a distributive lattice with respect to $\cup$ and $\cap$ as lattice operations.

$\Updownarrow$

There exists a partially ordered set (poset) $\mathcal{P} = (\Pi(E), \preceq)$ on a partition $\Pi(E)$ of E such that $\mathcal{D}$ is given by

$$\mathcal{D} = \{X \subseteq E \mid \exists \text{ **ideal** } \mathcal{J} \text{ of } \mathcal{P}\text{: } X = \bigcup_{F \in \mathcal{J}} F\}.$$

($X \subseteq E$ is an **ideal** of $\mathcal{P}$ if $e \in X$ and $e' \preceq e$ always imply $e' \in X$.)



(a) $\mathcal{D}$ (b) $\mathcal{P}(\mathcal{D})$

Theorem (Birkhoff-Iri): $\mathcal{D} \subseteq 2^E$, $\emptyset, E \in \mathcal{D}$.

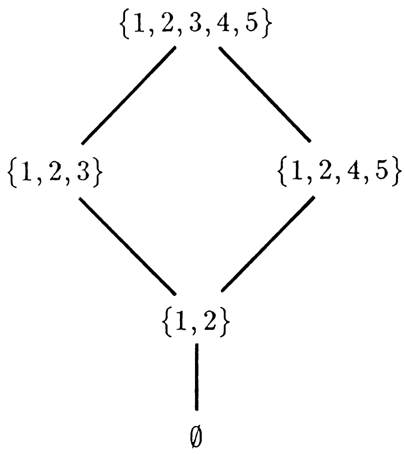
$\mathcal{D}$ is a distributive lattice with respect to $\cup$ and $\cap$ as lattice operations.

$\Updownarrow$

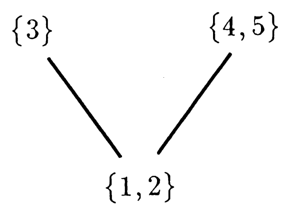
There exists a partially ordered set (poset) $\mathcal{P} = (\Pi(E), \preceq)$ on a partition $\Pi(E)$ of E such that $\mathcal{D}$ is given by

$$\mathcal{D} = \{X \subseteq E \mid \exists \text{ ideal } \mathcal{J} \text{ of } \mathcal{P}\colon X = \bigcup_{F \in \mathcal{J}} F\}.$$

($X \subseteq E$ is an **ideal** of $\mathcal{P}$ if $e \in X$ and $e' \preceq e$ always imply $e' \in X$.)



(a) $\mathcal{D}$



(b) $\mathcal{P}(\mathcal{D})$

A **simple distributive lattice** $\mathcal{D}$: every component of partition $\Pi(E)$ is a singleton, i.e., $\mathcal{D} = \mathbf{2}^{\mathcal{P}}$ (the set of all ideals of $\mathcal{P}=(E, \preceq)$).

$\longrightarrow$

Submodular System $(\mathcal{D}, f)$ on E

$\mathcal{D} \subseteq 2^E$: **a distributive lattice** $\quad (\emptyset, E \in \mathcal{D})$

$$X,\ Y \in \mathcal{D} \implies X \cup Y,\ X \cap Y \in \mathcal{D}$$

$f(: \mathcal{D} \to \mathbf{R})$: **a submodular function** $\quad (f(\emptyset) = 0)$

$$\forall X, Y \in \mathcal{D} : f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

Submodular System $(\mathcal{D}, f)$ on E

$\mathcal{D} \subseteq 2^E$: **a distributive lattice** $(\emptyset, E \in \mathcal{D})$

$$X,\ Y \in \mathcal{D} \implies X \cup Y,\ X \cap Y \in \mathcal{D}$$

$f(: \mathcal{D} \to \mathbf{R})$: **a submodular function** $(f(\emptyset) = 0)$

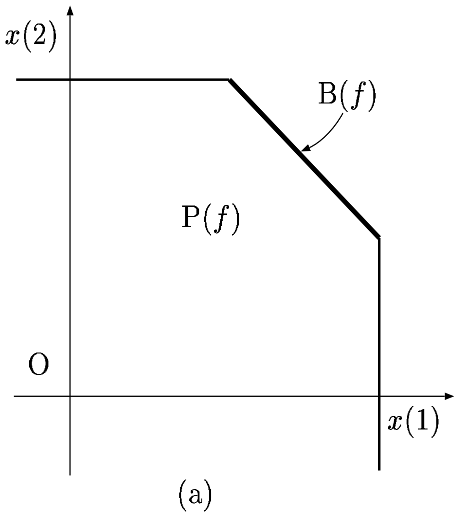
$$\forall X, Y \in \mathcal{D} : f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

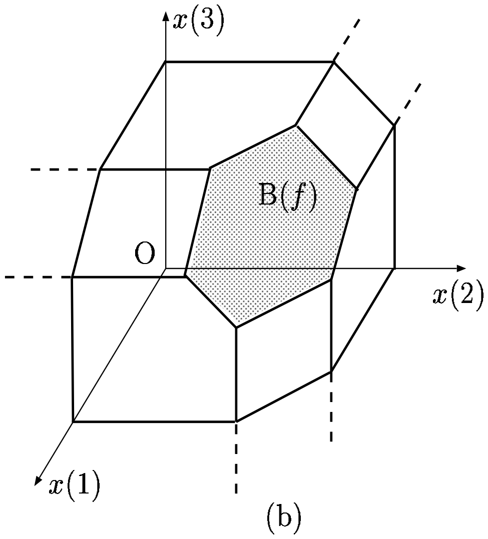
$$\mathrm{P}(f) = \{x \in \mathbf{R}^E \mid \forall X \in \mathcal{D} : x(X) \leq f(X)\}$$

(**Submodular Polyhedron**)

$$x(X) = \sum_{e \in X} x(e), \qquad x(\emptyset) = 0$$

$\mathrm{B}(f) = \{x \mid x \in \mathrm{P}(f),\ x(E) = f(E)\}$ (**Base Polyhedron**)





$\mathcal{D} \subseteq 2^E$: **a distributive lattice** $(\emptyset, E \in \mathcal{D})$

$$X,\ Y \in \mathcal{D} \implies X \cup Y,\ X \cap Y \in \mathcal{D}$$

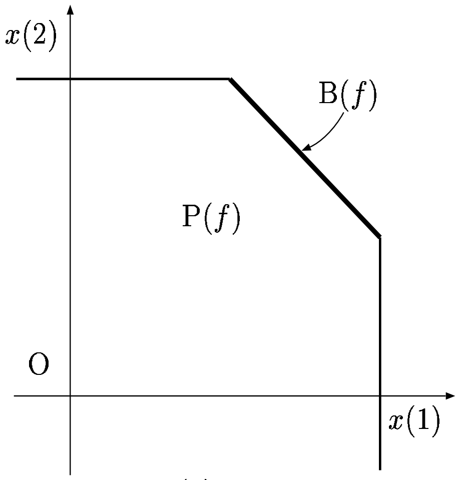
$f(: \mathcal{D} \to \mathbf{R})$: **a submodular function** $(f(\emptyset) = 0)$

$$\forall X, Y \in \mathcal{D} : f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

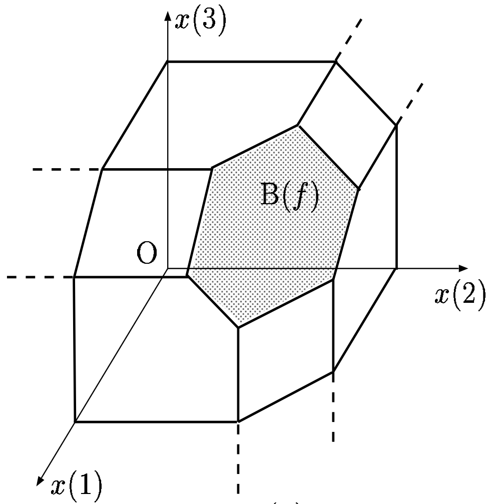
$\mathrm{P}(f) = \{x \in \mathbf{R}^E \mid \forall X \in \mathcal{D} : x(X) \leq f(X)\}$
(**Submodular Polyhedron**)

$$x(X) = \sum_{e \in X} x(e), \qquad x(\emptyset) = 0$$

$\mathrm{B}(f) = \{x \mid x \in \mathrm{P}(f),\ x(E) = f(E)\}$ (**Base Polyhedron**)



(a)



(b)

Remark: **Submodular system** $(\mathcal{D}, f) \overset{1:1}{\longleftrightarrow}$ **Base polyhedron** $\mathrm{B}(f)$
(**Submodular polyhedron** $\mathrm{P}(f)$)

$\longrightarrow$

Define a **supermodular system** $(\mathcal{D}, g)$ and its associated **supermodular polyhedron** $\mathrm{P}(g)$ and **base polyhedron** $\mathrm{B}(g)$ in a dual manner.

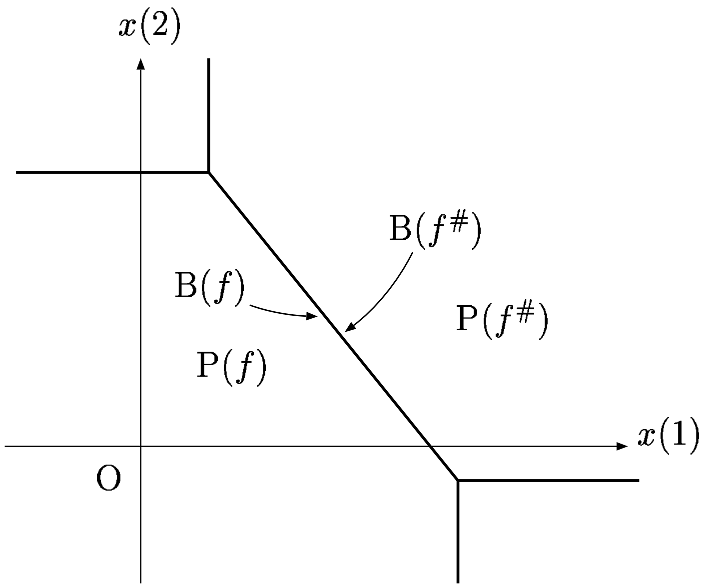
Define a **supermodular system** $(\mathcal{D}, g)$ and its associated **supermodular polyhedron** $\mathrm{P}(g)$ and **base polyhedron** $\mathrm{B}(g)$ in a dual manner.

Duality

$$\bar{\mathcal{D}} = \{E \setminus X \mid X \in \mathcal{D}\}$$

$$f^{\#}(E \setminus X) = f(E) - f(X) \qquad (X \in \mathcal{D})$$

$(\bar{\mathcal{D}}, f^{\#})$: **the supermodular system dual to submodular system** $(\mathcal{D}, f)$

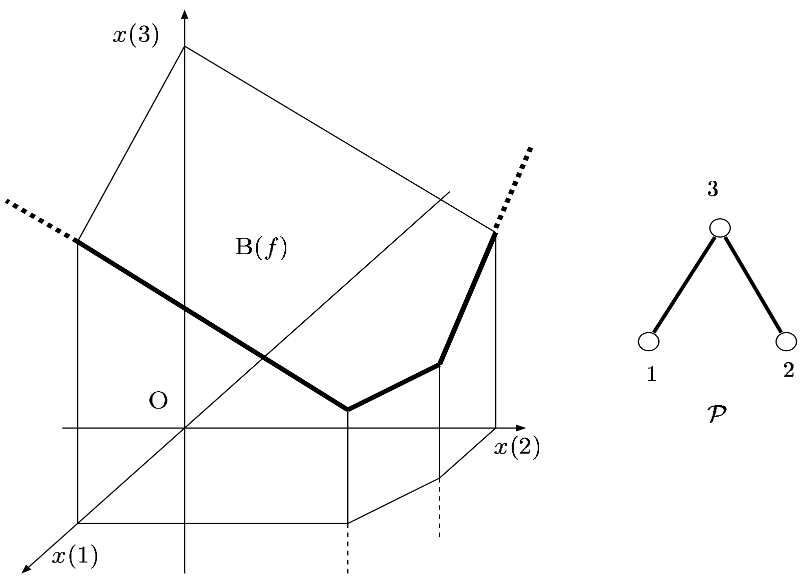


$$\mathrm{B}(f) = \mathrm{B}(f^{\#})$$

$\longrightarrow$

$(\mathcal{D}, f)$: **A submodular system on** E

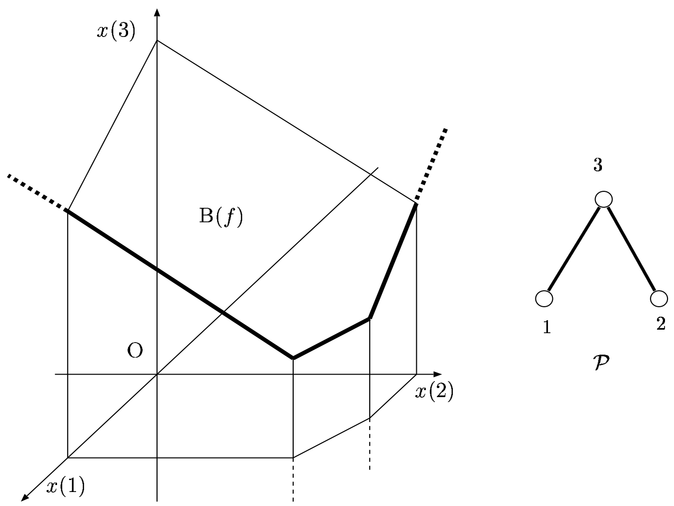
Proposition: The base polyhedron of $(\mathcal{D}, f)$ has an **extreme point**.
$\iff$ $\mathcal{D}$ is **simple**.



$(\mathcal{D}, f)$: **A submodular system on** E

Proposition: The base polyhedron of $(\mathcal{D}, f)$ has an extreme point.
$\iff$ $\mathcal{D}$ is simple.

Proposition: The base polyhedron of $(\mathcal{D}, f)$ is **bounded**.
$\iff$ $\mathcal{D} = 2^E$ (a **Boolean** lattice)

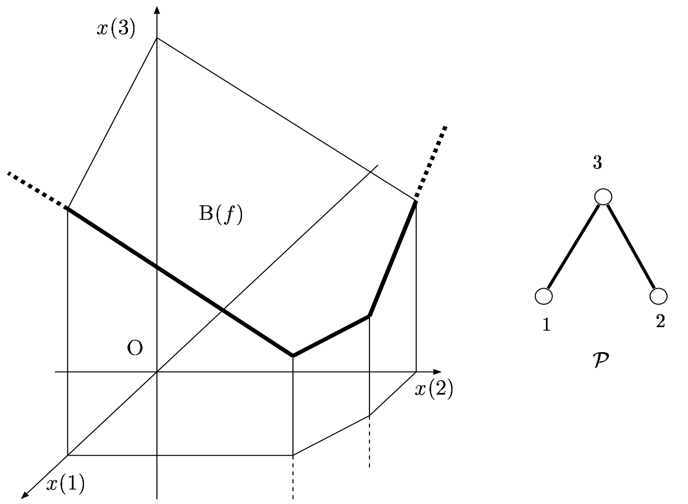


$(\mathcal{D}, f)$: **A submodular system on** E

Proposition: The base polyhedron of $(\mathcal{D}, f)$ has an extreme point.
$\iff$ $\mathcal{D}$ is simple.

Proposition: The base polyhedron of $(\mathcal{D}, f)$ is bounded.
$\iff$ $\mathcal{D} = 2^E$ (a Boolean lattice)

Proposition: Suppose $\mathcal{D}$ is simple. Then,
all **the extreme bases are nonnegative**
$\iff$ f is **monotone nondecreasing**.



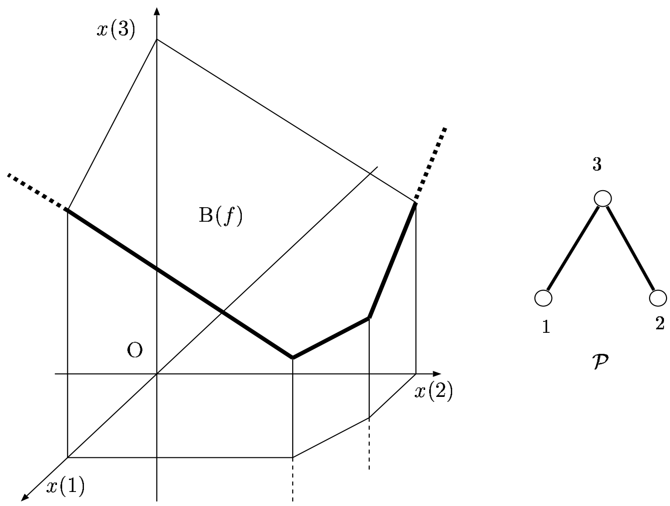
$(\mathcal{D}, f)$: **A submodular system on** E

Proposition: The base polyhedron of $(\mathcal{D}, f)$ has an extreme point.
$\iff$ $\mathcal{D}$ is simple.

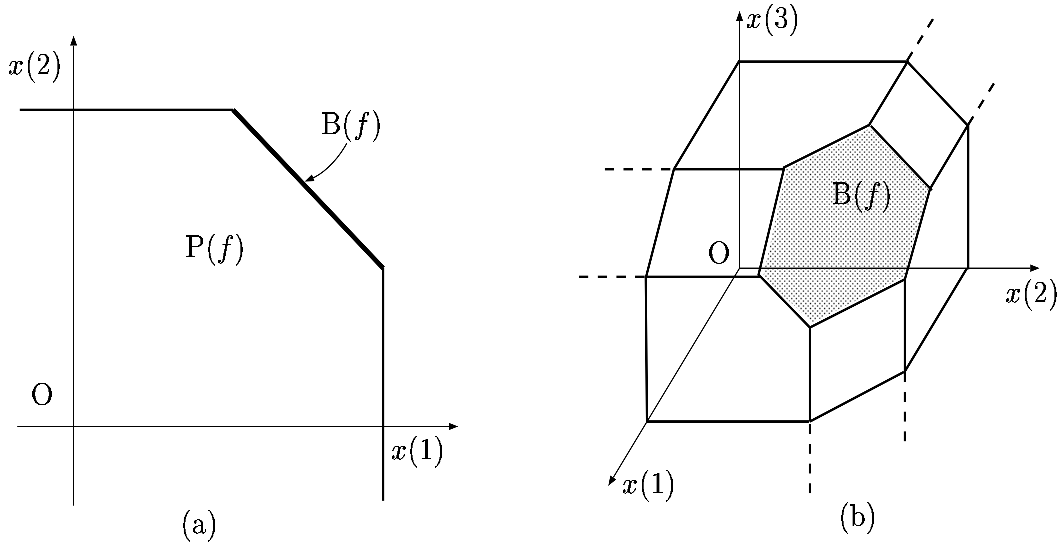
Proposition: The base polyhedron of $(\mathcal{D}, f)$ is bounded.
$\iff$ $\mathcal{D} = 2^E$ (a Boolean lattice)

Proposition: Suppose $\mathcal{D}$ is simple. Then,
all the extreme bases are nonnegative
$\iff$ f is monotone nondecreasing.

Proposition: Suppose $\mathcal{D}$ is simple. Then,
all **the extreme bases are integral** $\iff$ f is **integer-valued**.



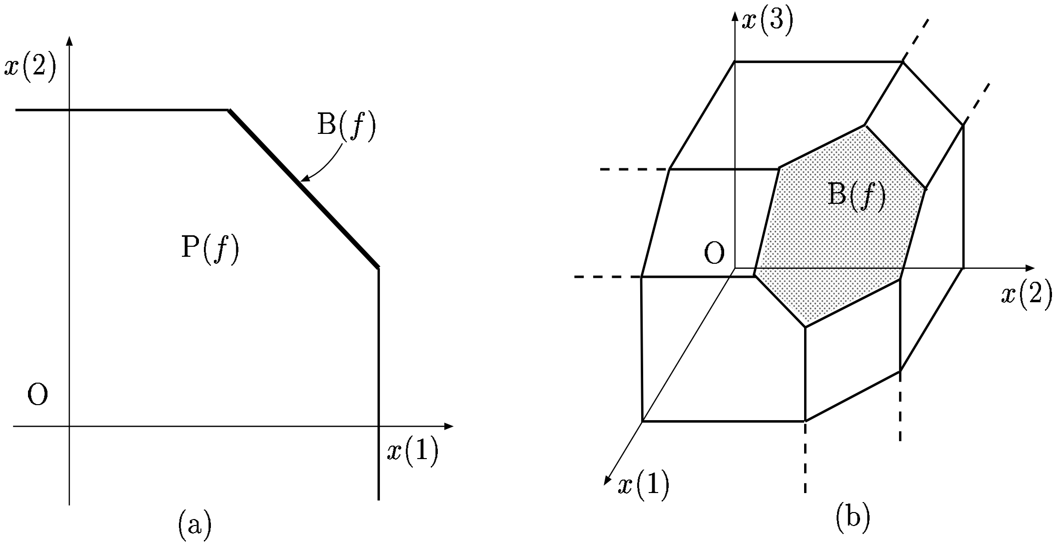
$\longrightarrow$



Polymatroid (Edmonds):

$\mathcal{D} = 2^E$ and f is monotone nondecreasing
$(X \subseteq Y \subseteq E \Longrightarrow f(X) \leq f(Y))$.

$\Longleftrightarrow \quad \mathrm{B}(f) \subset \mathbf{R}_+^E$



Polymatroid (Edmonds):

$\mathcal{D} = 2^E$ and f is monotone nondecreasing
$(X \subseteq Y \subseteq E \Longrightarrow f(X) \leq f(Y))$.

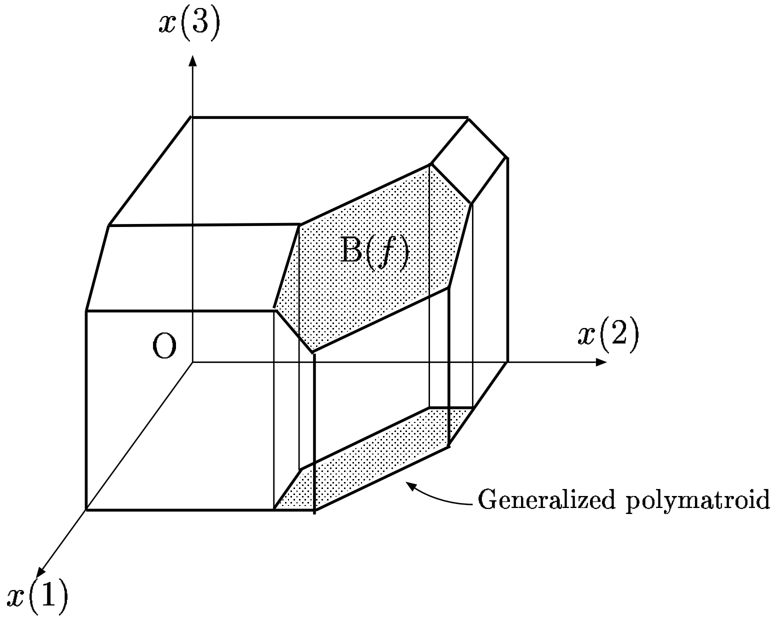
$\Longleftrightarrow \quad \mathrm{B}(f) \subset \mathbf{R}^E_+$

Matroid (Whitney):

Furthermore, f is integer-valued and has a unit-increase property.
$\forall X \in 2^E, \, \forall e \in E \setminus X : \; f(X) \leq f(X \cup \{e\}) \leq f(X) + 1$
(extreme bases $\longleftrightarrow$ matroid bases)

$\longrightarrow$

Generalized polymatroids (Frank, Hassin) and **Base Polyhedra**



$\longrightarrow$

Theorem (Tomizawa): For a bounded polyhedron $P \subset \mathbf{R}^E$,

P is a base polyhedron

$\Updownarrow$

all the edge vectors of P are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$

Theorem (Tomizawa): For a bounded polyhedron $P \subset \mathbf{R}^E$,

P is a base polyhedron

$\Updownarrow$

all the edge vectors of P are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$

Corollary: For a bounded polyhedron $P \subset \mathbf{R}^E$,

P is a generalized polymatroid

$\Updownarrow$

all the edge vectors of P are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$ or $(0, \cdots, 0, \pm 1, 0, \cdots, 0)$

Remark: The above two are also valid for pointed polyhedra.

$\longrightarrow$

The Intersection Theorem and Its Equivalents

$(\mathcal{D}_i, f_i)$ $(i = 1, 2)$: submodular systems on E

The Intersection Theorem (Edmonds):

$$\max\{x(E) \mid x \in \mathrm{P}(f_1) \cap \mathrm{P}(f_2)\}$$
$$= \min\{f_1(X) + f_2(E \setminus X) \mid X \in \mathcal{D}_1,\ E \setminus X \in \mathcal{D}_2\}$$

(+ Integrality)

The Intersection Theorem and Its Equivalents

$(\mathcal{D}_i, f_i)$ $(i = 1, 2)$: submodular systems on E

The Intersection Theorem (Edmonds):

$$\max\{x(E) \mid x \in \mathrm{P}(f_1) \cap \mathrm{P}(f_2)\}$$
$$= \min\{f_1(X) + f_2(E \setminus X) \mid X \in \mathcal{D}_1,\ E \setminus X \in \mathcal{D}_2\}$$

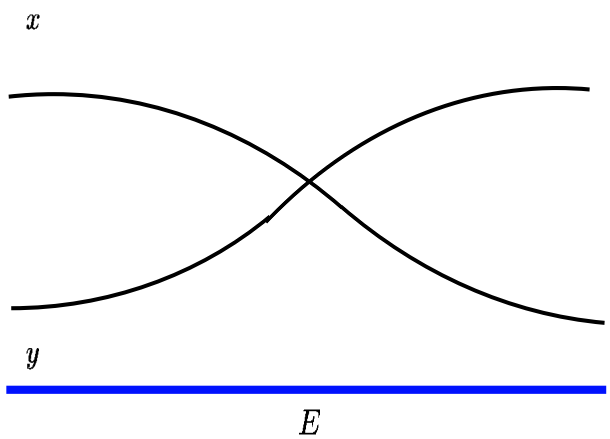
(+ Integrality)

The Intersection Theorem' :

$$\max\{x \wedge y(E) \mid x \in \mathrm{B}(f_1),\ \ y \in \mathrm{B}(f_2)\}$$
$$= \min\{f_1(X) + f_2(E \setminus X) \mid X \in \mathcal{D}_1,\ E \setminus X \in \mathcal{D}_2\}$$

(+ Integrality)

$$(x \wedge y)(e) = \min\{x(e), y(e)\} \ \ (e \in E).$$



$\longrightarrow$

$(\mathcal{D}_1, f)$, $(\mathcal{D}_2, g)$: a submodular system and supermodular system on E

Discrete Separation Theorem (Frank):

$$f \geq g \quad \Longrightarrow \quad \exists z \in \mathbf{R}^E : f \geq z \geq g \quad (\text{i.e., } \mathrm{P}(f) \cap \mathrm{P}(g) \neq \emptyset)$$

(+ Integrality)

$(\mathcal{D}_1, f)$, $(\mathcal{D}_2, g)$: a submodular system and supermodular system on E

Discrete Separation Theorem (Frank):

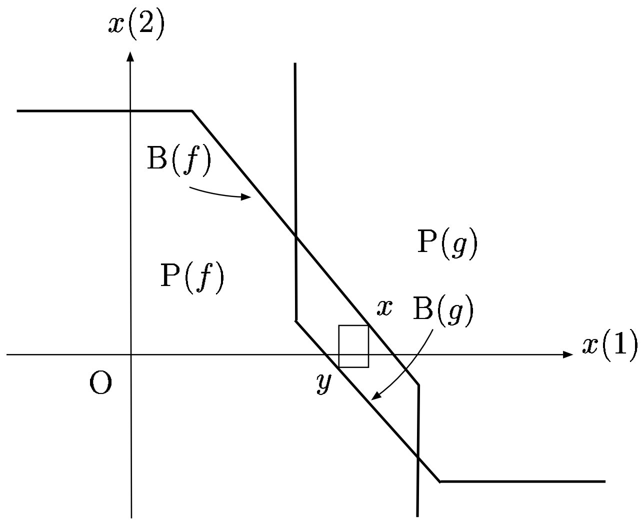
$$f \geq g \quad \Longrightarrow \quad \exists z \in \mathbf{R}^E : f \geq z \geq g \quad (\text{i.e., } \mathrm{P}(f) \cap \mathrm{P}(g) \neq \emptyset)$$

(+ Integrality)

Discrete Separation Theorem':

$$f \geq g \quad \Longrightarrow \quad \exists (x \in \mathrm{B}(f),\ y \in \mathrm{B}(g)) : x \geq y$$

(+ Integrality)



$\longrightarrow$

$(\mathcal{D}_1, f)$: a submodular system on E
$(\mathcal{D}_2, g)$: a supermodular system on E

$$f^*(x) = \max\{x(X) - f(X) \mid X \in \mathcal{D}_1\} \quad (x \in \mathbf{R}^E)$$
$$g^*(x) = \min\{x(X) - g(X) \mid X \in \mathcal{D}_2\} \quad (x \in \mathbf{R}^E)$$

Fenchel Duality Theorem (F):

$$\min\{f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2\}$$
$$= \max\{g^*(x) - f^*(x) \mid x \in \mathbf{R}^E\}$$

(+ Integrality)

$(\mathcal{D}_1, f)$: a submodular system on E
$(\mathcal{D}_2, g)$: a supermodular system on E

$$f^*(x) = \max\{x(X) - f(X) \mid X \in \mathcal{D}_1\} \quad (x \in \mathbf{R}^E)$$
$$g^*(x) = \min\{x(X) - g(X) \mid X \in \mathcal{D}_2\} \quad (x \in \mathbf{R}^E)$$

Fenchel Duality Theorem (F):

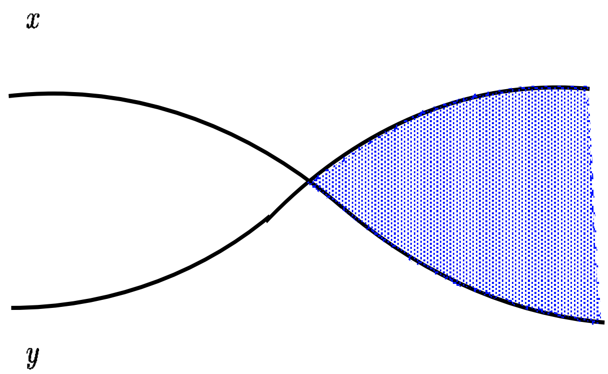
$$\min\{f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2\}$$
$$= \max\{g^*(x) - f^*(x) \mid x \in \mathbf{R}^E\}$$

(+ Integrality)

Fenchel Duality Theorem':

$$\min\{f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2\}$$
$$= \max\{(x-y)^-(E) \mid x \in \mathrm{B}(f),\ y \in \mathrm{B}(g)\}$$

(+ Integrality) $((x-y)^- = (\min\{0, x(e) - y(e)\} \mid e \in E))$



$\longrightarrow$

$(\mathcal{D}_i, f_i)$ $(i = 1, 2)$: submodular systems on E

Minkowski Sum Theorem:

$$\mathrm{P}(f_1) + \mathrm{P}(f_2) = \mathrm{P}(f_1 + f_2),$$
$$\mathrm{B}(f_1) + \mathrm{B}(f_2) = \mathrm{B}(f_1 + f_2).$$

Moreover, if f_1 and f_2 are integer-valued, the collections $\mathrm{P}_{\mathbf{Z}}(\cdot)$ and $\mathrm{B}_{\mathbf{Z}}(\cdot)$ of integer points in $\mathrm{P}(\cdot)$ and $\mathrm{B}(\cdot)$ satisfy

$$\mathrm{P}_{\mathbf{Z}}(f_1) + \mathrm{P}_{\mathbf{Z}}(f_2) = \mathrm{P}_{\mathbf{Z}}(f_1 + f_2),$$
$$\mathrm{B}_{\mathbf{Z}}(f_1) + \mathrm{B}_{\mathbf{Z}}(f_2) = \mathrm{B}_{\mathbf{Z}}(f_1 + f_2).$$

Minimum-Norm Base and SFM

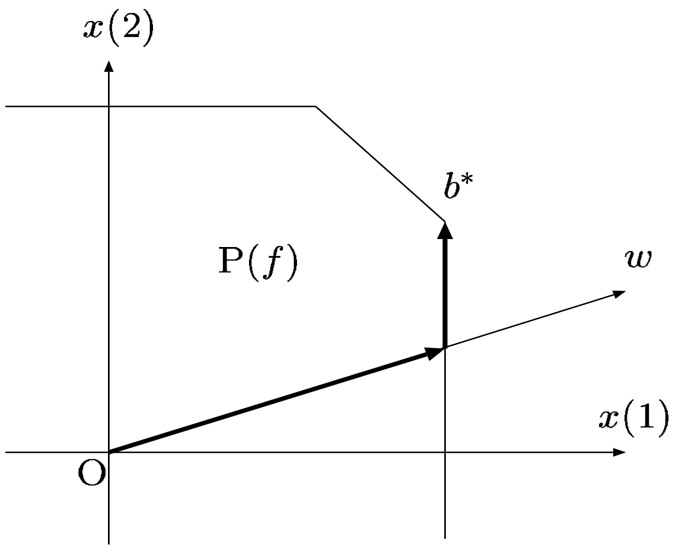
$(\mathcal{D}, f)$: a submodular system on E
$w(: E \to \mathbf{R})$: a positive weight function, $\quad \lambda \in \mathbf{R}$

Minimum-Norm Base and SFM

$(\mathcal{D}, f)$: a submodular system on E
$w(: E \to \mathbf{R})$: a positive weight function, $\quad \lambda \in \mathbf{R}$

Parametric Vector Reduction: For any $\lambda \in \mathbf{R}$,

$$\max\{x(E) \mid x \in \mathrm{P}(f),\ x \le \lambda w\}$$
$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}$$



$\longrightarrow$

Theorem (F): There exists a **unique base** b^* such that for all $\lambda \in \mathbf{R}$ $x = b^* \wedge \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in \mathrm{P}(f),\ x \leq \lambda w\}$$
$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$

Theorem (F): There exists a **unique base** b^* such that for all $\lambda \in \mathbf{R}$ $x = b^* \wedge \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in \mathrm{P}(f),\ x \leq \lambda w\}$$
$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$

A base $\hat{b}$ is called a **lexicographically optimal base** w.r.t. weight w if it **lexicographically maximizes** the sequence

$$T(b/w) = (b(e_1)/w(e_1), \cdots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e)$ $(e \in E)$ arranged in nondecreasing order of magnitude among all bases b.

Theorem (F): There exists a **unique base** b^* such that for all $\lambda \in \mathbf{R}$ $x = b^* \wedge \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in \mathrm{P}(f),\ x \leq \lambda w\}$$
$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$

A base $\hat{b}$ is called a **lexicographically optimal base** w.r.t. weight w if it **lexicographically maximizes** the sequence

$$T(b/w) = (b(e_1)/w(e_1), \cdots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e)$ $(e \in E)$ arranged in nondecreasing order of magnitude among all bases b.

Theorem (F):

(1) b^* is the **lexicographically optimal base** with respect to weight w.

(2) b^* is the **minimizer of** $\sum_{e \in E} b^2(e)/w(e)$ **over** $\mathrm{B}(f)$**.**

$\longrightarrow$

Theorem (F): There exists a **unique base** b^* such that for all $\lambda \in \mathbf{R}$ $x = b^* \wedge \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in \mathrm{P}(f),\ x \leq \lambda w\}$$
$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$

A base $\hat{b}$ is called a **lexicographically optimal base** w.r.t. weight w if it **lexicographically maximizes** the sequence

$$T(b/w) = (b(e_1)/w(e_1), \cdots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e)$ $(e \in E)$ arranged in nondecreasing order of magnitude among all bases b.

Theorem (F):

(1) b^* is the **lexicographically optimal base** with respect to weight w.

(2) b^* is the **minimizer of** $\sum_{e \in E} b^2(e)/w(e)$ **over** $\mathrm{B}(f)$.

(⟶ Resource Allocation Problems + submodular constraints)

→

Remarks: We have

$$\min\{f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D}\}$$
$$= (b^* \wedge \lambda w)(E) = \max\{x(E) \mid x \leq \lambda w,\ x \in \mathrm{P}(f)\}$$

Hence, for $\lambda = 0$

$$\min\{f(X) \mid X \in \mathcal{D}\} = \max\{x(E) \mid x \leq \mathbf{0},\ x \in \mathrm{P}(f)\},$$

Remarks: We have

$$\min\{f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D}\}$$
$$= (b^* \wedge \lambda w)(E) = \max\{x(E) \mid x \leq \lambda w,\ x \in \mathrm{P}(f)\}$$

Hence, for $\lambda = 0$

$$\min\{f(X) \mid X \in \mathcal{D}\} = \max\{x(E) \mid x \leq \mathbf{0},\ x \in \mathrm{P}(f)\},$$

which implies

$$A_0 = \{e \mid b^*(e) \leq 0\}, \qquad A_- = \{e \mid b^*(e) < 0\}$$

are the **unique maximal** and the **unique minimal minimizer** of f.

Remarks: We have

$$\min\{f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D}\}$$
$$= (b^* \wedge \lambda w)(E) = \max\{x(E) \mid x \leq \lambda w,\ x \in \mathrm{P}(f)\}$$

Hence, for $\lambda = 0$

$$\min\{f(X) \mid X \in \mathcal{D}\} = \max\{x(E) \mid x \leq \mathbf{0},\ x \in \mathrm{P}(f)\},$$

which implies

$$A_0 = \{e \mid b^*(e) \leq 0\}, \qquad A_- = \{e \mid b^*(e) < 0\}$$

are the **unique maximal** and the **unique minimal minimizer** of f.

---

b^* is the minimizer of $\sum_{e \in E} b^2(e)/w(e)$

Minimum-norm base

$\Longrightarrow$ **Submodular Function Minimization**

Applicability of P. Wolfe's minimum-norm point algorithm

$\longrightarrow$

(Submodular Function Minimization $\longleftarrow$ MNP algorithm of Wolfe)

$$A_0 = \{e \mid b^*(e) \le 0\}, \qquad A_- = \{e \mid b^*(e) < 0\}$$

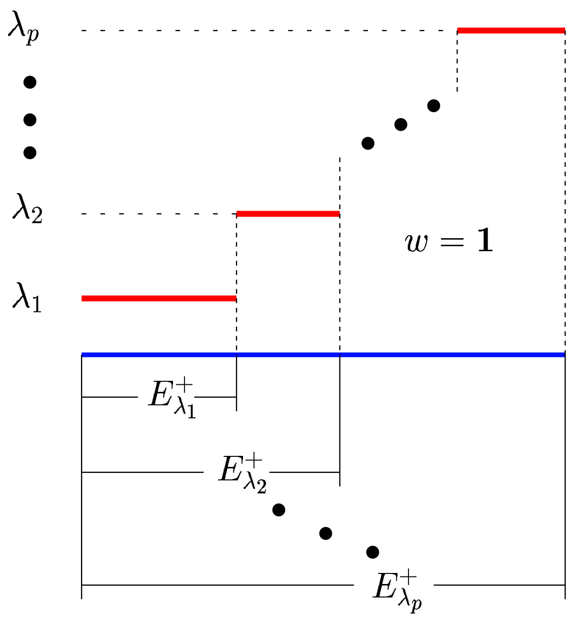
When f is **integer-valued and** $w = \mathbf{1}$,

$$\lambda_i = \frac{f(E^+_{\lambda_i}) - f(E^+_{\lambda_{i-1}})}{|E^+_{\lambda_i} \setminus E^+_{\lambda_{i-1}}|} \quad (i = 1, 2, \cdots, p), \qquad E^+_{\lambda_0} = \emptyset$$

$$\lambda_1 < \cdots < \lambda_p$$

$$\min_i\{\lambda_i > 0\} \ge \frac{1}{|E|}, \quad \max_i\{\lambda_i < 0\} \le -\frac{1}{|E|}$$

$$A_0 = \{e \mid b^*(e) \le \epsilon\}, \qquad A_- = \{e \mid b^*(e) < -\epsilon\} \quad (\epsilon = \tfrac{1}{2|E|})$$



$\longrightarrow$

$\mathcal{D} = 2^E$

Maximum Weight Base Problem

A weight function $w : E \to [0, 1]$

$$\begin{array}{ll} \text{Maximize} & \displaystyle\sum_{e \in E} w(e)x(e) \\ \text{subject to} & x \in \mathrm{B}(f) \end{array}$$

$\mathcal{D} = 2^E$

Maximum Weight Base Problem

A weight function $w : E \to [0, 1]$

$$\text{Maximize} \quad \sum_{e \in E} w(e)x(e)$$
$$\text{subject to} \quad x \in \mathrm{B}(f)$$

For a permutation $\sigma = (e_1, e_2, \cdots, e_n)$ of E, define

$$\Delta(\sigma): \quad 1 \geq x(e_1) \geq x(e_2) \geq \cdots \geq x(e_n) \geq 0$$

$$S_i = \{e_1, \cdots, e_i\} \quad (i = 1, \cdots, n)$$

$$S_0 = \emptyset \subset S_1 \subset \cdots \subset S_n = E$$

and

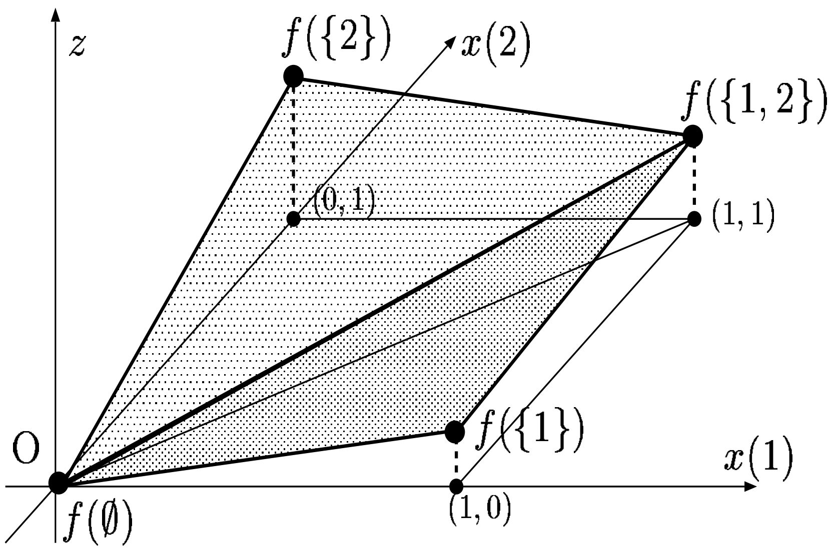
$$b^\sigma(e_i) = f(S_i) - f(S_{i-1}) \quad (i = 1, \cdots, n)$$

Then, b^σ is a **maximum weight base** for $w \in \Delta(\sigma)$.
($\longleftarrow$ **Greedy Algorithm**) (Edmonds)

$\longrightarrow$

$\hat{f}(w)$ = the **value of a maximum weight base**
(the **support function** of $\mathrm{B}(f)$ restricted on $[0,1]^E$)

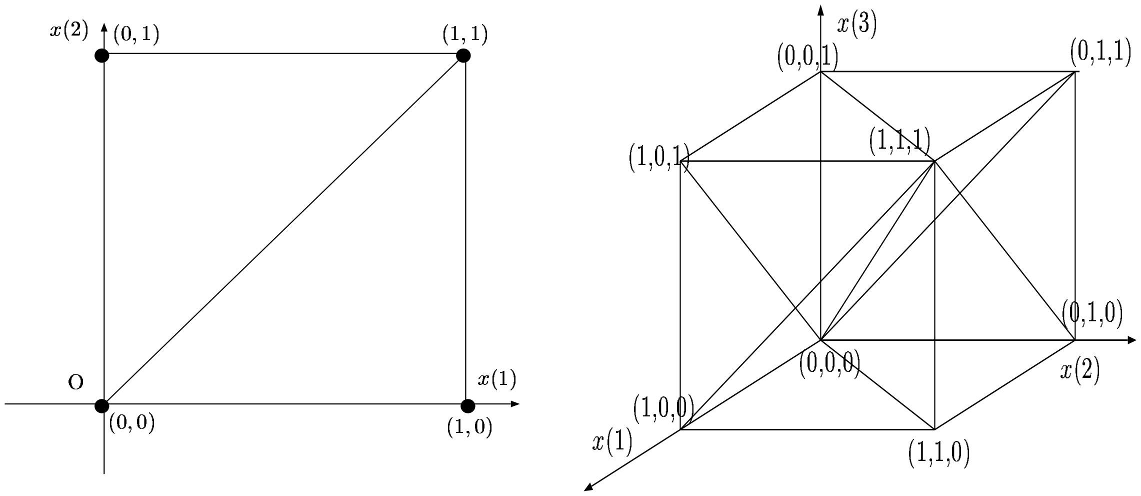
(*) $\hat{f}$ is a **linear function on each cell** $\Delta(\sigma)$.



$\longrightarrow$

$$\Delta(\sigma): \quad 1 \geq x(e_1) \geq x(e_2) \geq \cdots \geq x(e_n) \geq 0$$

The set of cells for $n!$ permutations σ defines a **simplicial division** of unit hypercube $[0,1]^E$ (**Freudenthal simplicial division**).



For **any set function** $f: 2^E \to \mathbf{R}$, the **piecewise-linear function** $\hat{f}$ obtained by linear interpolation on every cell $\Delta(\sigma)$ is called the **Lovász extension** (or the Choquet integral) of f.

$\longrightarrow$

Theorem (Lovász): For any set function $f : 2^E \to \mathbf{R}$, f is a **submodular** function if and only if its **Lovász extension $\hat{f}$ is convex**.

Theorem (Lovász): For any set function $f : 2^E \to \mathbf{R}$, f is a submodular function if and only if its Lovász extension $\hat{f}$ is convex.

In other words,

Submodular functions
$\Longleftrightarrow$ **Convex extensible** w.r.t. the **Freudenthal simplicial division**

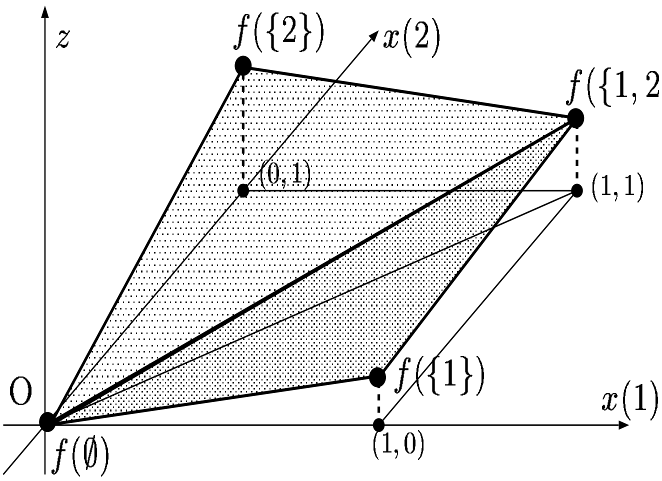
$\longrightarrow$

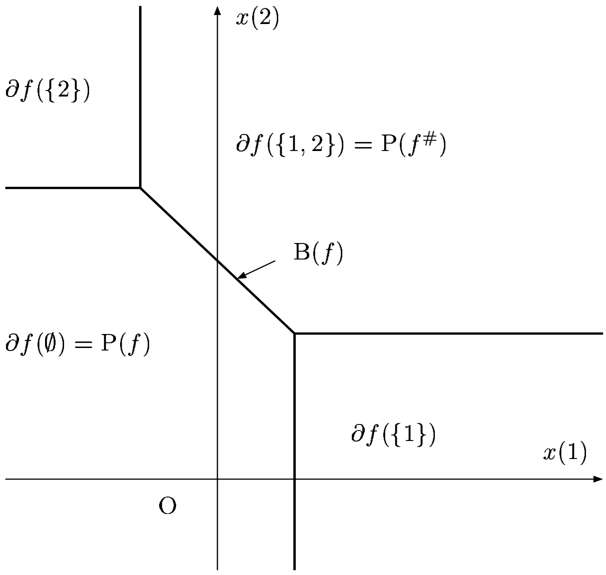
Subgradients and Subdifferentials of Submodular Functions

The **subdifferential** of f at $X \in \mathcal{D}$

$$\partial f(X) = \{x \in \mathbf{R}^E \mid \forall Y \in \mathcal{D} : x(Y) - x(X) \leq f(Y) - f(X)\}$$

Note: $\langle x, \chi_Y - \chi_X \rangle \leq f(\chi_Y) - f(\chi_X), \quad \partial f(X) = \partial \hat{f}(\chi_X)$





$\longrightarrow$

Submodular function $f : 2^E \to \mathbf{R}$

$\Updownarrow$

The Lovász extension $\hat{f}$ is convex (and linear on every cell $\Delta(\sigma)$)

$\Updownarrow$

Base polyhedron $\mathrm{B}(f)$ (edge vectors $(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$)

$\Updownarrow$

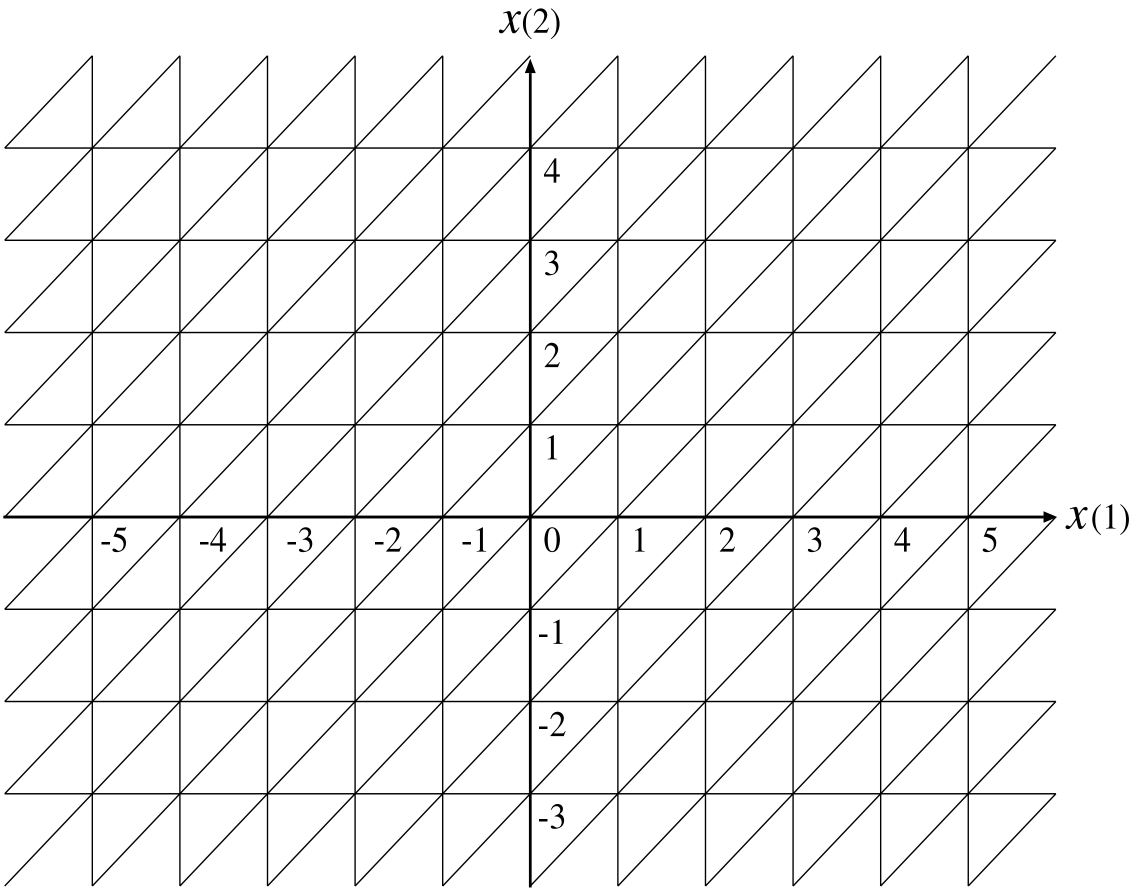
Greedy algorithm works

$\longrightarrow$

Submodular function $f : 2^E \to \mathbf{R}$		Base polyhedron $\mathrm{B}(f)$
$\Updownarrow$		$\Updownarrow$
Lovász extension $\hat{f}$ is convex		Greedy algorithm
$\Downarrow$		$\Downarrow$
Submodular integrally convex function		Valuated matroid
(Favati-Tardella (1990))		(Dress-Wenzel (1990, 1992))
$\Updownarrow$	(Convex conjugate)	$\Downarrow$
L-/L$^\natural$-convex function	$\longleftrightarrow$	M-/M$^\natural$-convex function
(Murota (1998), F-Murota(2000))		(Murota(1996), Murota-Shioura (1999))

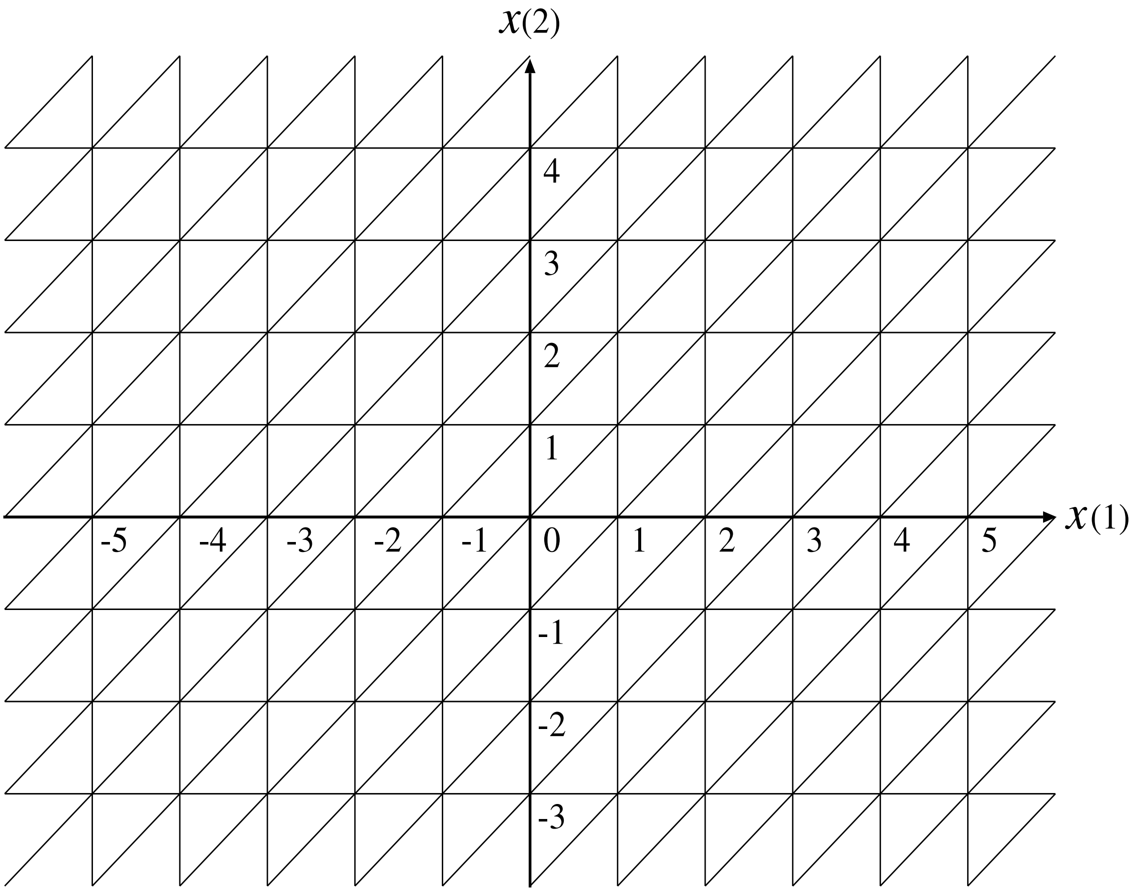
Discrete Convex Analysis (Kazuo Murota)

A simplicial division of the plane (triangulation)



The Freudenthal simplicial division

A simplicial division of the plane (triangulation)



The Freudenthal simplicial division

Consider a function f defined on the integer lattice $\mathbf{Z}^n$.

$\longrightarrow$

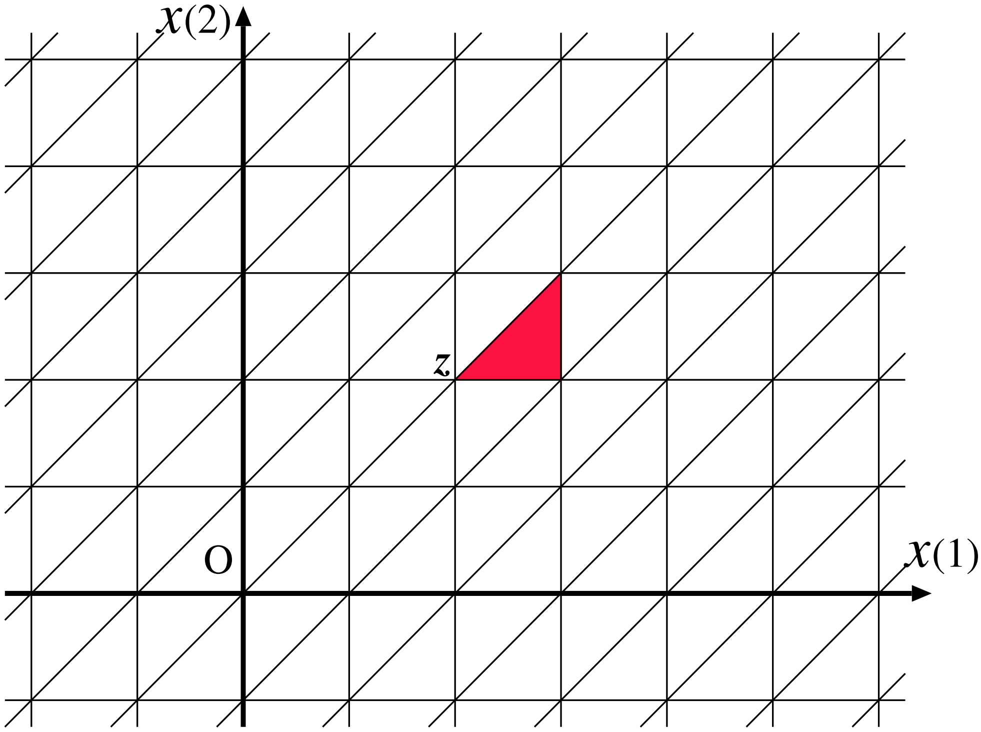
Discrete convex functions with respect to **Freudenthal simplicial division** = **L$^\natural$-convex functions** defined on $\mathbf{Z}^n$ (due to Murota)

This is equivalent to the

Submodular integrally convex function

due to Favati and Tardella (1990)

(**L$^\natural$-concave functions** are defined similarly.)

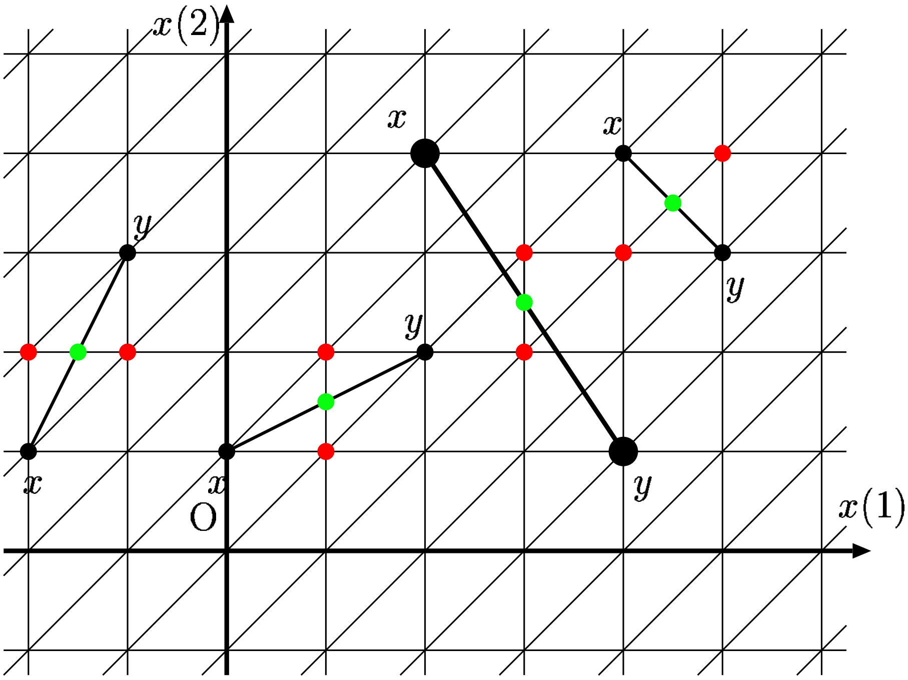


The Freudenthal simplicial division of $\mathbf{R}^2$.

$\longrightarrow$

Characterization by **mid-point convexity** due to Favati-Tardella

$$f(x) + f(y) \geq f(\lceil \tfrac{1}{2}(x+y) \rceil) + f(\lfloor \tfrac{1}{2}(x+y) \rfloor) \qquad (\forall x,\ y \in \mathbf{Z}^n).$$



Remark:

$$x + y = \lceil \tfrac{1}{2}(x+y) \rceil + \lfloor \tfrac{1}{2}(x+y) \rfloor,$$

$$\hat{f}(\tfrac{1}{2}(x+y)) = \tfrac{1}{2}\{f(\lceil \tfrac{1}{2}(x+y) \rceil) + f(\lfloor \tfrac{1}{2}(x+y) \rfloor)\}$$

$$\Longrightarrow \quad \tfrac{1}{2}\{\hat{f}(x) + \hat{f}(y)\} \geq \hat{f}(\tfrac{1}{2}(x+y)) \qquad (\forall x,\ y \in \mathbf{Z}^n).$$

$\longrightarrow$

f: **$\mathrm{L}^\natural$-convex function** on integer lattice $\mathbf{Z}^n$

$\hat{f}$: the convex extension of f on the Freudenthal simplicial division

---

Convex conjugate $\hat{f}^\bullet$ of $\hat{f}$ (or f) (**Legendre-Fenchel transform**)

$$\begin{aligned}\hat{f}^\bullet(p) &= \sup\{\langle p, x\rangle - \hat{f}(x) \mid x \in \mathbf{R}^n\} \\ &= \sup\{\langle p, x\rangle - f(x) \mid x \in \mathbf{Z}^n\} \qquad (p \in (\mathbf{R}^n)^*)\end{aligned}$$

where $\langle p, x\rangle = \sum_{i=1}^{n} p(i)x(i)$.

$\hat{f}^\bullet(p)$ is an **$\mathrm{M}^\natural$-convex function** (Murota-Shioura)

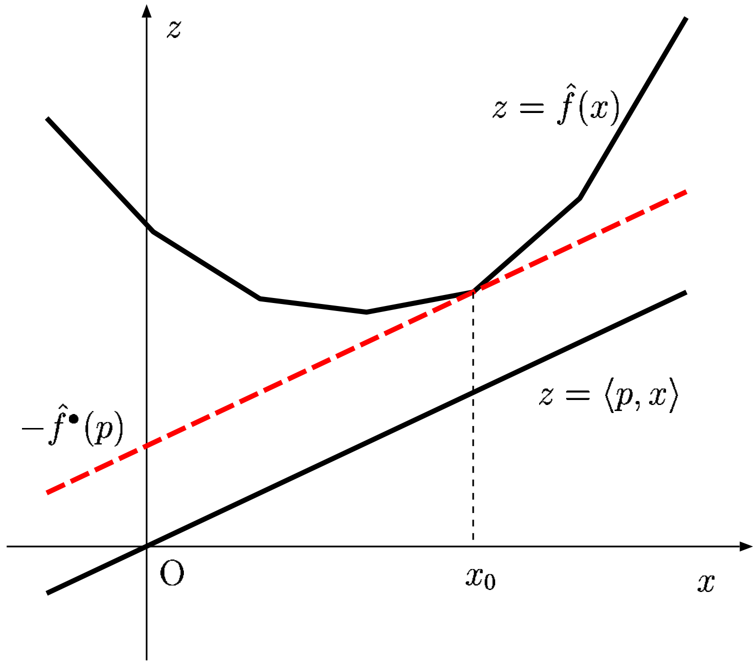
$f^\bullet$: restriction of convex conjugate $\hat{f}^\bullet$ on $(\mathbf{Z}^n)^*$

f is integer-valued $\Longrightarrow$ $\hat{f}^\bullet$ is the convex extension of $f^\bullet$

$f^\bullet$ for integer-valued f is exactly an integer-valued $\mathrm{M}^\natural$-convex function on $(\mathbf{Z}^n)^*$ (Murota-Shioura)

$\longrightarrow$

$$\begin{aligned}\hat{f}^{\bullet}(p) &= \sup\{\langle p, x\rangle - \hat{f}(x) \mid x \in \mathbf{R}^n\} \\ &= \sup\{\langle p, x\rangle - f(x) \mid x \in \mathbf{Z}^n\} \qquad (p \in (\mathbf{R}^n)^*)\end{aligned}$$



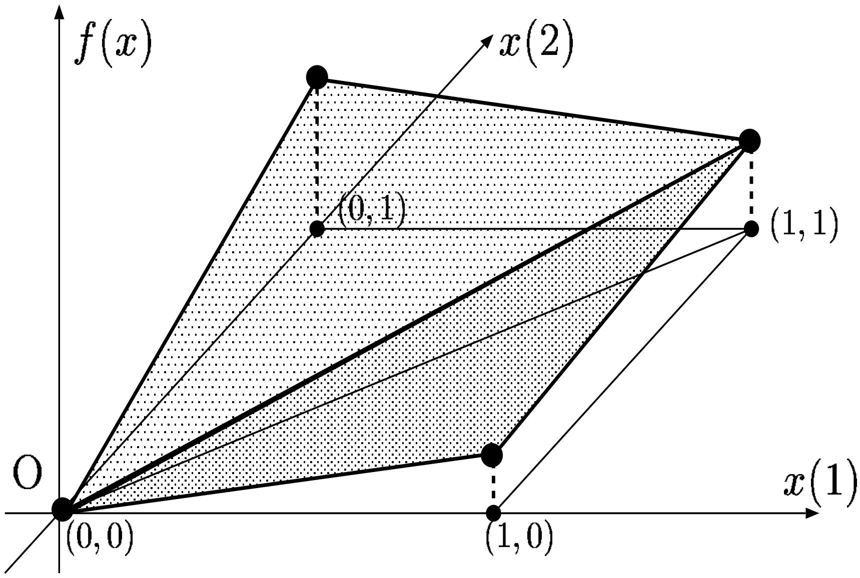
For $x_0 \in \mathbf{Z}^n$ the set of all $p \in (\mathbf{R}^n)^*$ satisfying

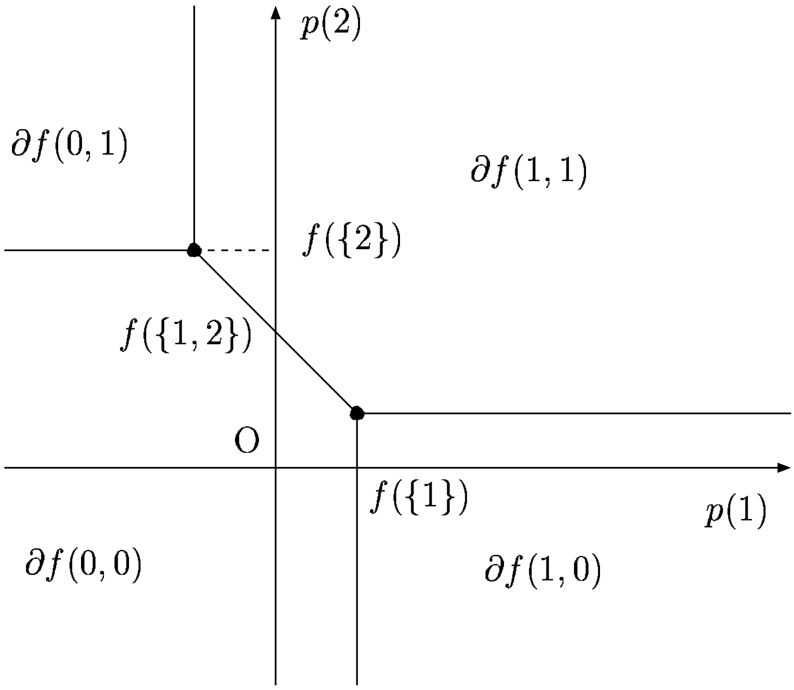
$$\hat{f}^{\bullet}(p) = \langle p, x_0\rangle - f(x_0), \quad \text{i.e.}$$

$$f(x) \geq f(x_0) + \langle p, x - x_0\rangle \quad (\forall x \in \mathbf{Z}^n)$$

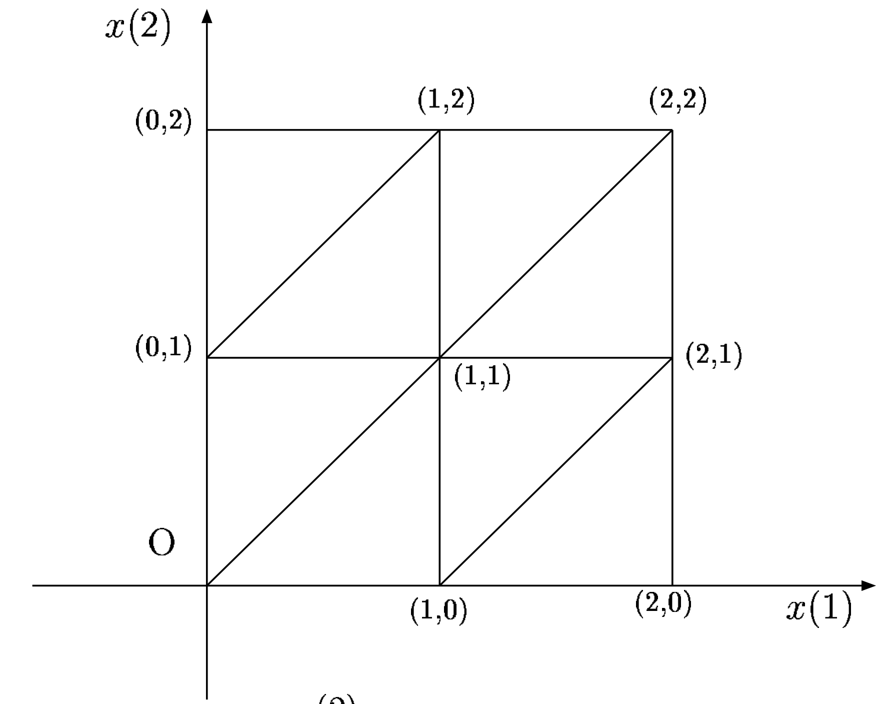
is the **subdifferential** $\partial \hat{f}(x_0) (= \partial f(x_0))$ **at** x_0.

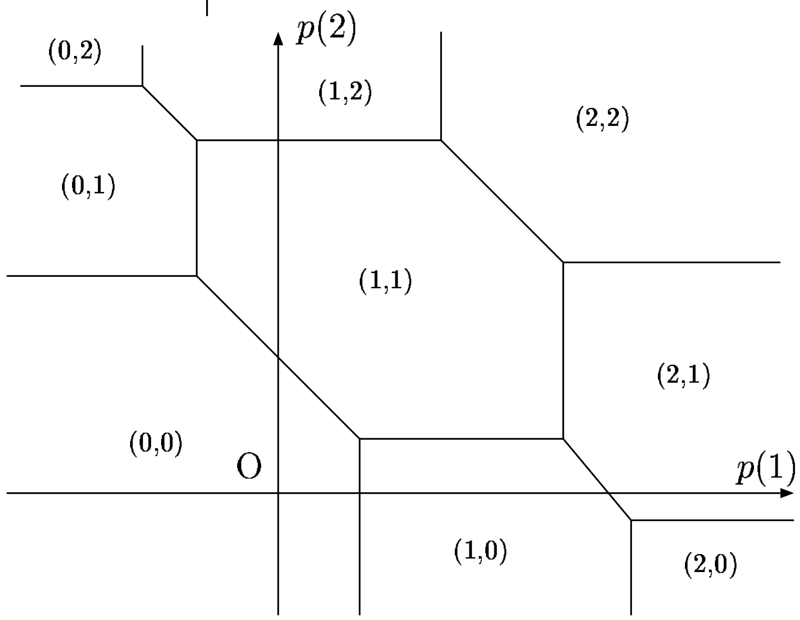
$\longrightarrow$





$\longrightarrow$





$\longrightarrow$

(Recall)

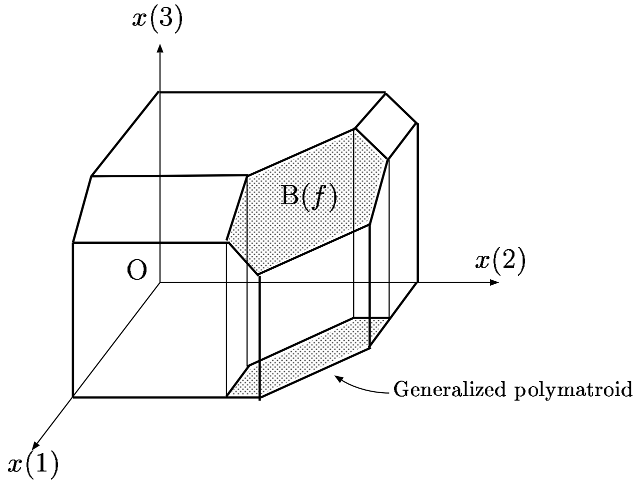
Corollary: For a pointed polyhedron $P \subset \mathbf{R}^E$,

P is a generalized polymatroid

$\Updownarrow$

all the edge vectors of P are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$ or $(0, \cdots, 0, \pm 1, 0, \cdots, 0)$



$\longrightarrow$

(Recall)

Corollary: For a pointed polyhedron $P \subset \mathbf{R}^E$,

P is a generalized polymatroid

$$\Updownarrow$$

all the edge vectors of P are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$ or $(0, \cdots, 0, \pm 1, 0, \cdots, 0)$

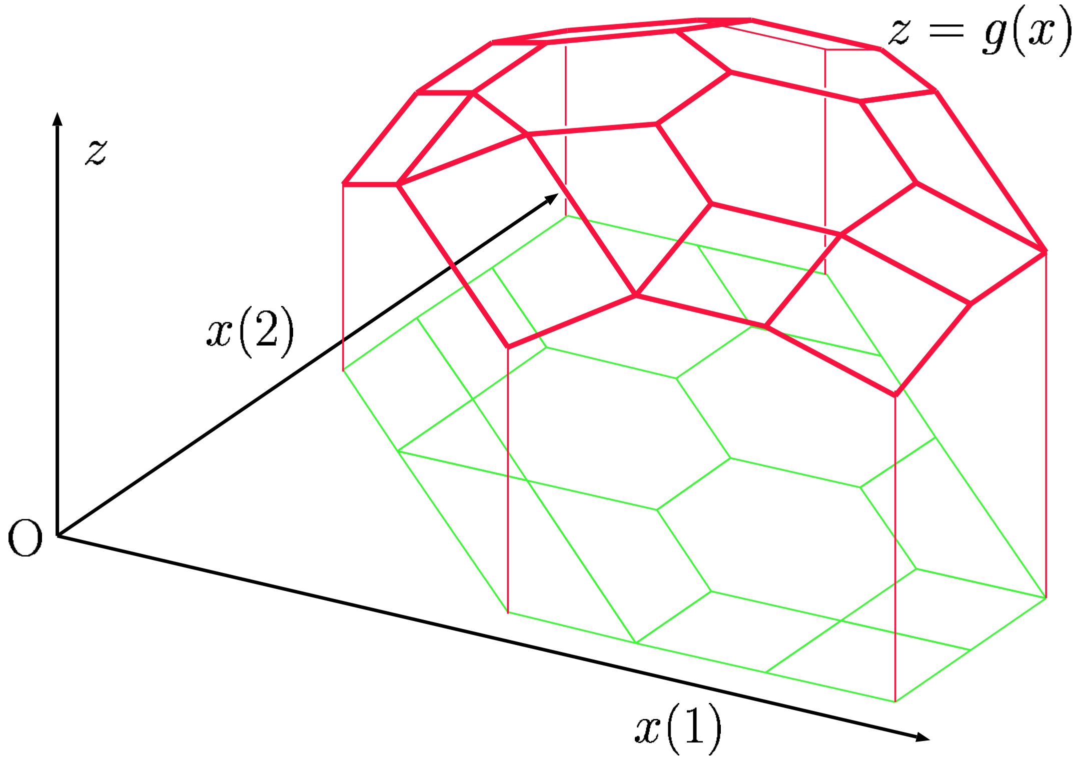
Hence, the **subdifferential** $\partial f(z)$ of an L$^\natural$-convex function f at an integral point z is a **generalized polymatroid**.

M$^\natural$-convex function $\hat{f}^\bullet$ is an **affine function** on every such **generalized polymatroid**.

Remark: If f is an **integer-valued** function, $\partial f(z)$ is an **integral** generalized polymatroid.

$\longrightarrow$

M$^\natural$-concave function g



$\longrightarrow$

Simultaneous Exchange Axiom for M$^\natural$-convex functions

$f : \mathbf{Z}^n \to \mathbf{R} \cup \{+\infty\}$ (Murota-Shioura)

$$\mathrm{dom}\, f = \{x \mid f(x) < +\infty\}$$

$$\mathrm{supp}^+(x) = \{i \mid x(i) > 0\}, \qquad \mathrm{supp}^-(x) = \{i \mid x(i) < 0\}$$

(M$^\natural$-EXC) For $x, y \in \mathrm{dom}\, f$ and $i \in \mathrm{supp}^+(x-y)$,

$$f(x) + f(y) \geq \min \Big[f(x - \chi_i) + f(y + \chi_i), \min_{j \in \mathrm{supp}^-(x-y)} \{ f(x - \chi_i + \chi_j) + f(y + \chi_i - \chi_j) \} \Big].$$

($\longrightarrow$ Simultaneous Exchange Axiom for **Generalized Polymatroids**)

$\longrightarrow$

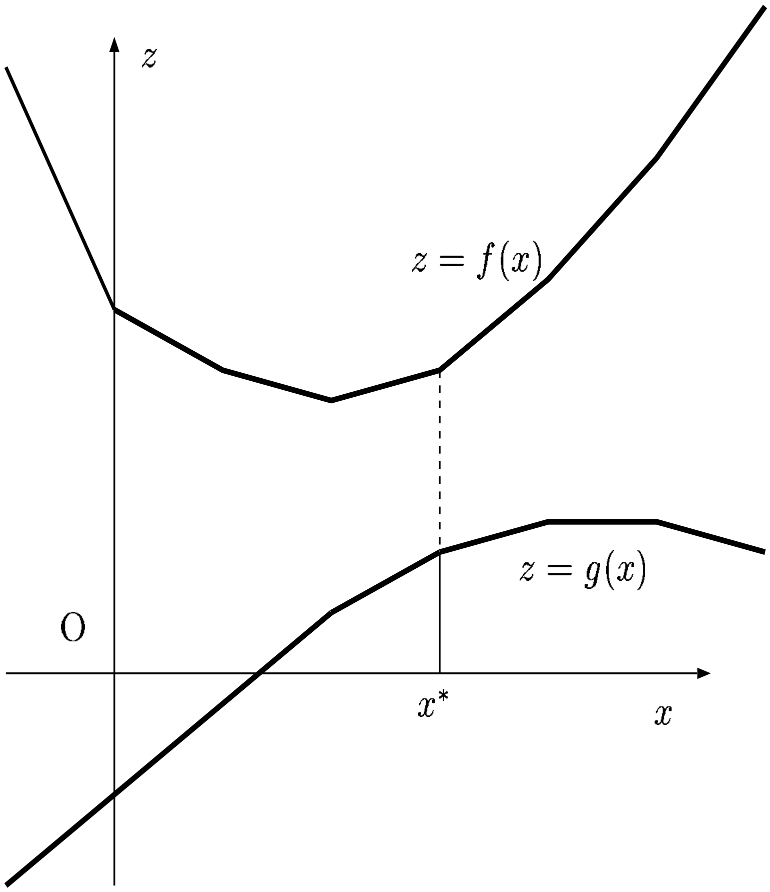
Discrete Separation Theorem ($L^\natural$)

f: an integer-valued $L^\natural$-convex function on $\mathbf{Z}^n$

g: an integer-valued $L^\natural$-concave function on $\mathbf{Z}^n$

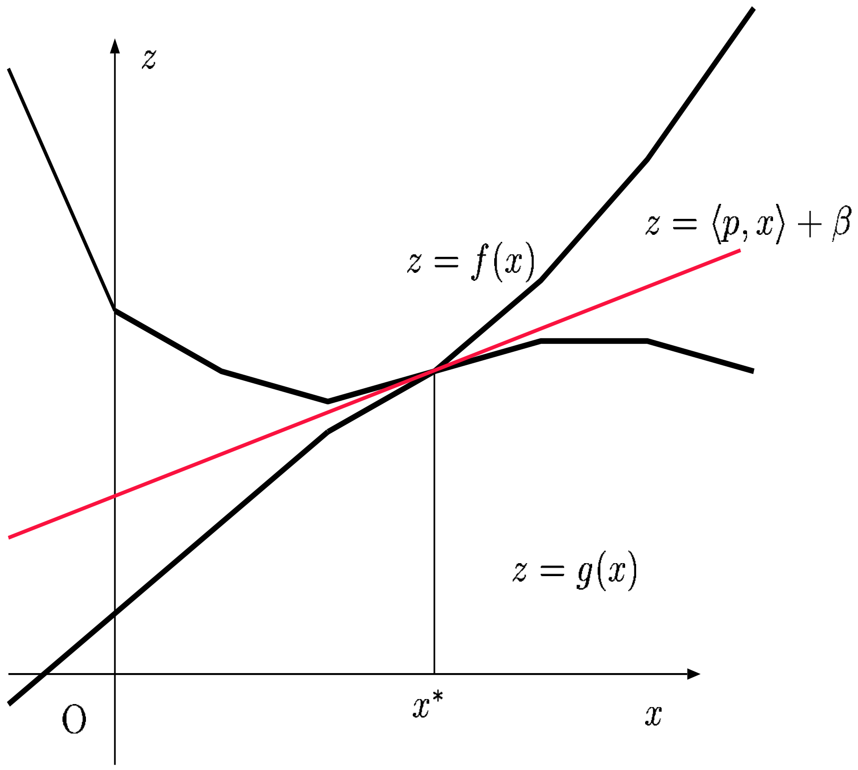
Suppose $f \geq g$.

$\longrightarrow$



Suppose x^* minimizes $f(x) - g(x)$ over $\mathbf{Z}^n$.

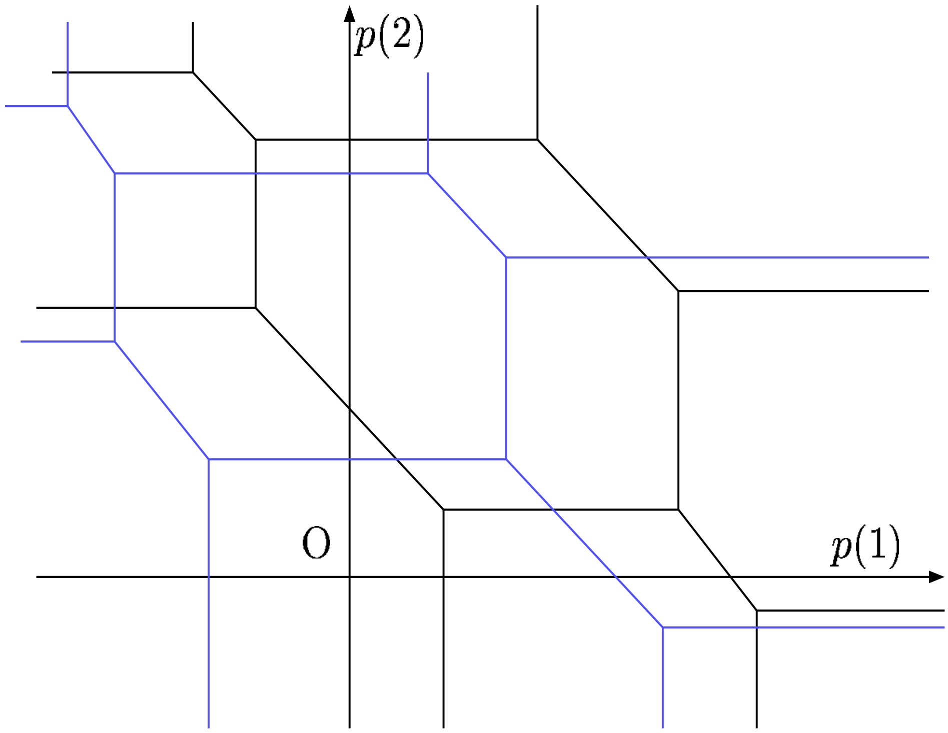
$\longrightarrow$



Common sub- and supergradient $p \in \partial f(x^*) \cap \partial g(x^*)$

$\longrightarrow$

Intersections of subdifferentials and superdifferentials (integral generalized polymatroids)



$\longrightarrow$

There exists a **common integral sub- and supergradient**

$$p \in \partial f(x^*) \cap \partial g(x^*)$$

due to the following integrality:

Intersection Theorem for Generalized Polymatroids (Edmonds, Frank):
The **intersection of any two integral generalized polymatroids is integral** if it is nonempty.

There exists a **common integral sub- and supergradient**

$$p \in \partial f(x^*) \cap \partial g(x^*)$$

due to the following integrality:

Intersection Theorem for Generalized Polymatroids (Edmonds, Frank): The **intersection of any two integral generalized polymatroids is integral** if it is nonempty.

Hence,

$$g(x^*) - \langle p, x^* \rangle \leq \beta \leq f(x^*) - \langle p, x^* \rangle$$

(p, x^*: **integral vectors $\Longrightarrow$ There exists an integer β.**

There exists a **common integral sub- and supergradient**

$$p \in \partial f(x^*) \cap \partial g(x^*)$$

due to the following integrality:

Intersection Theorem for Generalized Polymatroids (Edmonds, Frank):
The **intersection of any two integral generalized polymatroids is integral** if it is nonempty.

Hence,

$$g(x^*) - \langle p, x^* \rangle \leq \beta \leq f(x^*) - \langle p, x^* \rangle$$

(p, x^*: **integral vectors** $\Longrightarrow$ **There exists an integer** β**.**

Discrete Separation Theorem (Murota):

$\forall z \in \mathbf{Z}^n : f(z) \geq g(z)$

$$\Longrightarrow \quad \exists (p \in (\mathbf{Z}^n)^*, \ \beta \in \mathbf{Z}) : f(z) \geq \langle p, z \rangle + \beta \geq g(z) \ \ (\forall z \in \mathbf{Z}^n)$$

$\longrightarrow$

Discrete Fenchel Duality Theorem (Murota):

For any integer-valued L$^\natural$-convex function $f : \mathbf{Z}^n \to \mathbf{Z}$ and L$^\natural$-concave function $g : \mathbf{Z}^n \to \mathbf{Z}$,

$$\inf\{f(x) - g(x) \mid x \in \mathbf{Z}^n\} = \sup\{g^\circ(p) - f^\bullet(p) \mid p \in (\mathbf{Z}^n)^*\}.$$

Discrete Fenchel Duality Theorem (Murota):

For any integer-valued L$^\natural$-convex function $f : \mathbf{Z}^n \to \mathbf{Z}$ and L$^\natural$-concave function $g : \mathbf{Z}^n \to \mathbf{Z}$,

$$\inf\{f(x) - g(x) \mid x \in \mathbf{Z}^n\} = \sup\{g^\circ(p) - f^\bullet(p) \mid p \in (\mathbf{Z}^n)^*\}.$$

---

	conv. extension		**conv. conjugate**		**restriction**	
f	$\underset{\longleftarrow}{\longrightarrow}$	$\hat{f}$	$\longleftrightarrow$	$\hat{f}^\bullet$	$\underset{\longleftarrow}{\longrightarrow}$	$f^\bullet$
	restriction				**conv. extension**	
$\mathbf{Z}^n$		$\mathbf{R}^n$		$(\mathbf{R}^n)^*$		$(\mathbf{Z}^n)^*$

L$^\natural$-convex function **M$^\natural$-convex function**

For more information see the following monographs.

S. Fujishige: *Submodular Functions and Optimization*, Second Edition, Elsevier, 2005.

K. Murota: *Discrete Convex Analysis*, SIAM, 2003.

$\square$