# MATH 567 : Lecture 5 (01/23/2025)

Today: * representing sets
* representing functions

## Proof of Theorem 1 (continued..)

Recall...

**Theorem 1** $\bar{x}$ satisfies $(*)$ $\Longleftrightarrow$ $\exists\,(\bar{x}^1,\dots,\bar{x}^k, y_1,\dots,y_k)$ such that $(\bar{x}, \bar{x}^1,\dots,\bar{x}^k, y_1,\dots,y_k)$ satisfies $(*\text{-sharp})$.

$$\boxed{\begin{array}{c} A_1\bar{x}^1 \le \bar{b}^1 y_1 \\ \vdots \qquad \vdots \\ A_k\bar{x}^k \le \bar{b}^k y_k \\ \bar{x}^1+\bar{x}^2+\cdots+\bar{x}^k = \bar{x} \\ y_1+y_2+\cdots+y_k = 1 \\ y_i \in \{0,1\} \end{array}} \quad (*\text{-sharp})$$

Proof $(\Rightarrow)$: seen in the last lecture...

$(\Leftarrow)$ WLOG, let $y_1=1,\ y_i=0,\ i=2,\dots,k$ in $(\bar{x},\bar{x}^1,\dots,\bar{x}^k, y_1,\dots,y_k)$ that satisfies $(*\text{-sharp})$.

$$\Rightarrow \quad \begin{array}{l} A_1\bar{x}^1 \le \bar{b}^1 \\ A_2\bar{x}^2 \le \bar{0} \\ \quad\vdots \qquad \vdots \\ A_k\bar{x}^k \le \bar{0} \end{array} \quad \text{and} \quad \bar{x}^1+\cdots+\bar{x}^k = \bar{x}$$

from Assumption 2

$$\Rightarrow \quad A_1\bar{x} = A_1(\bar{x}^1+\cdots+\bar{x}^k) = A_1\bar{x}^1 + A_1\bar{x}^2+\cdots+A_1\bar{x}^k \le \bar{b}^1 + \bar{0}+\cdots+\bar{0}$$

$\Rightarrow A_1\bar{x} \le \bar{b}^1$, i.e., $\bar{x}$ satisfies $(*)$. $\square$

# Representing Sets in general

Q. In general, what all sets could we represent using 0-1 and/or general integer (GI) variables?

We need a few new definitions to address this question. In particular, we will formally define a formulation – so far, we have been studying them informally as MIP models.

**Def** A set $S$ is **bounded MIP-representable** (b-MIP-r) if $\exists$ matrices $A, B, C, D$ and a vector $\bar{f}$ such that

$$S = \{ (\bar{x}, \bar{y}) \in \mathbb{Z}^n \times \mathbb{R}^m \mid \exists (\bar{u}, \bar{v}) \in \mathbb{Z}^p \times \mathbb{R}^q \text{ such that } A\bar{x} + B\bar{y} + C\bar{u} + D\bar{v} \leq \bar{f} \}$$

and

$A\bar{x} + B\bar{y} + C\bar{u} + D\bar{v} \leq \bar{f}$ implies lower and upper bounds on $\bar{x}$ and $\bar{u}$ (the general integer (GI) variables). → hence the "bounded" in b-MIP-r

The set $P = \{ (\bar{x}, \bar{y}, \bar{u}, \bar{v}) \in \mathbb{R}^{n+m+p+q} \mid A\bar{x} + B\bar{y} + C\bar{u} + D\bar{v} \leq \bar{f} \}$ is called a **formulation** of $S$.

Note that all variables are continuous in this formal definition of a formulation.

**Def** $P = \{ \bar{x} \in \mathbb{R}^n \mid A\bar{x} \leq \bar{b} \}$ is a **polyhedron**, where $A \in \mathbb{R}^{m \times n}$, $\bar{b} \in \mathbb{R}^m$, $m, n$ are finite. → polyhedra are convex sets.

A bounded polyhedron is a **polytope**.

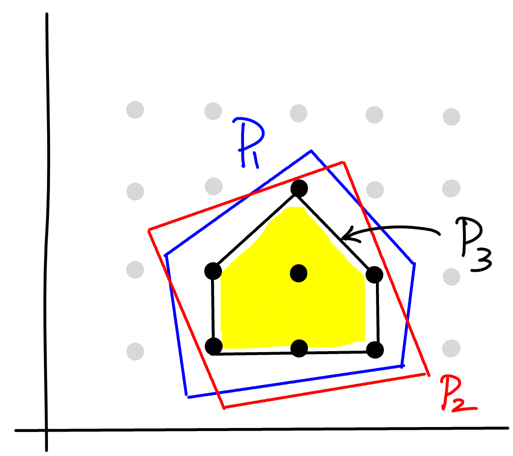
## Examples of formulations

1. $S = \{\bar{x} \in \{0,1\}^n \mid (x_1=1) \vee \cdots \vee (x_n=1)\}$ has a formulation

$$P = \{\bar{x} \in \mathbb{R}^n \mid 0 \le x_i \le 1 \ \forall i, \ x_1 + \cdots + x_n \ge 1\}.$$

2. $S = \{7$ lattice points shown as $\bullet\}$.

$P_1, P_2, P_3$ are all formulations for $S$.

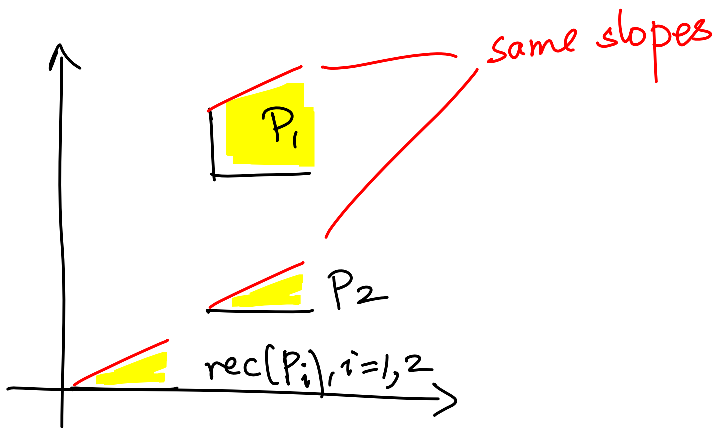



But $P_3$ is "better" than $P_1$ and $P_2$. Note that $P_3$ is the convex hull of the lattice points that is $S$. If we want to maximize a linear function over $S$, we could do the same over $P_3$ instead. The same claim cannot be made for $P_1$ or $P_2$. We will talk later about how to compare formulations.

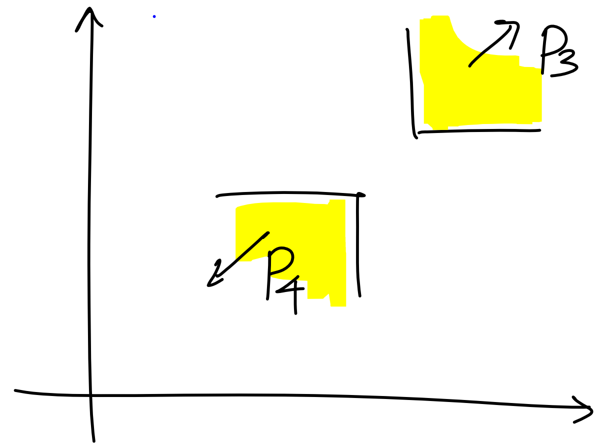
What kinds of sets $S$ are b-MIP-r? Could we characterize them?

**Theorem 2 (Jeraslow, Lowe):** $S$ is b-MIP-r iff $S = P_1 \cup \cdots \cup P_k$ for finite $k$, where $P_i$ are polyhedra having the same recession cone i.e., $\text{rec}(P_i)$ is independent of $i$, $i=1,\dots,k$.

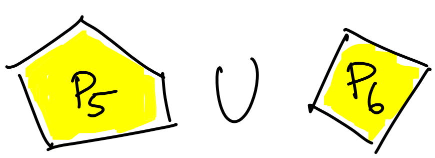
Here are some examples.



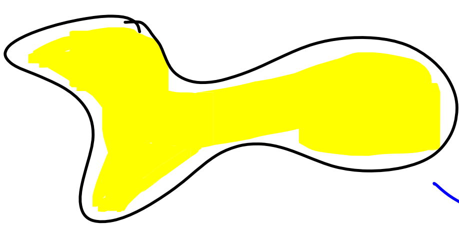

$P_1 \cup P_2$ is b-MIP-r.



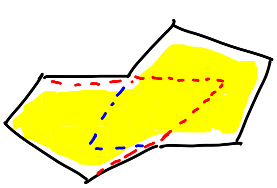

$P_3 \cup P_4$ is not okay as an MIP.
(as $\text{rec}(P_3) \neq \text{rec}(P_4)$)

$S =$  is okay as an MIP.


$(\text{rec}(P_5) = \text{rec}(P_6) = \{\bar{0}\})$

$S =$  is not okay as MIP, as it is not a polyhedron to start with.

→ $S$ being non convex is not crucial here — it could've been an ellipse, and the conclusion is the same.

$S =$  → o.k. as MIP! → union of two polytopes.

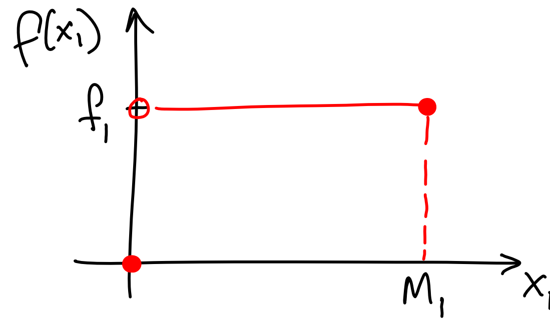
Can we use similar techniques to model functions, instead of sets?

# Representing Sets v/s Functions

We already saw the case of fixed charge:

$$\left\{\begin{array}{ll} \min & f(x_1) \\ \text{s.t.} & 0 \le x_1 \le M_1 \\ & A\bar{x} \le \bar{b} \end{array}\right\}$$

and we wrote an MIP for the same.




It turns out we could use similar ideas to model some classes of functions appearing in certain optimization problems. Naturally, if $f(x_1)$ is nonlinear, e.g., $x_1^3$ or $\sqrt{x_1}$, we will not get an integer linear program! We need some definitions first.

**Def** Given $f: \mathbb{R}^n \to \mathbb{R}$, we define the **graph**, **epigraph**, and **hypograph** of $f$ as follows.

$$\text{graph}(f) = \{(z, \bar{x}) \in \mathbb{R}^{n+1} \mid z = f(\bar{x})\},$$

$$\text{epi}(f) = \{(z, \bar{x}) \in \mathbb{R}^{n+1} \mid z \ge f(\bar{x})\}, \text{ and}$$

$$\text{hypo}(f) = \{(z, \bar{x}) \in \mathbb{R}^{n+1} \mid z \le f(\bar{x})\}.$$

Notice that graph$(f)$, epi$(f)$, and hypo$(f)$ are sets, and we could consider when each of them is b-MIP-r — instead of talking about representability of $f(\cdot)$ itself.

Suppose we have $\left\{ \begin{array}{l} \min f(\bar{x}) \\ \text{s.t. } A\bar{x} \le \bar{b} \end{array} \right\}$ where epi(f) is b-MIP-r.

Then we can write $\left\{ \begin{array}{l} \min\ z \\ \text{s.t. } A\bar{x} \le \bar{b} \\ \quad (z, \bar{x}) \in \text{epi}(f) \end{array} \right\}$ as the MIP representation.

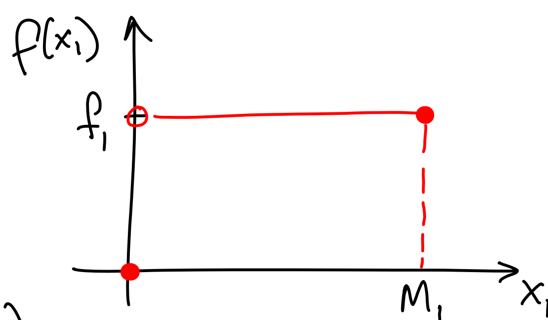
Since epi(f) is b-MIP-r, we can write down an MIP representation of epi(f), which completes the MIP model above.

Q. Why not require graph(f) being b-MIP-r?

**Theorem 3** graph(f) is b-MIP-r iff both epi(f) and hypo(f) are b-MIP-r.

## Example

The fixed charge function.

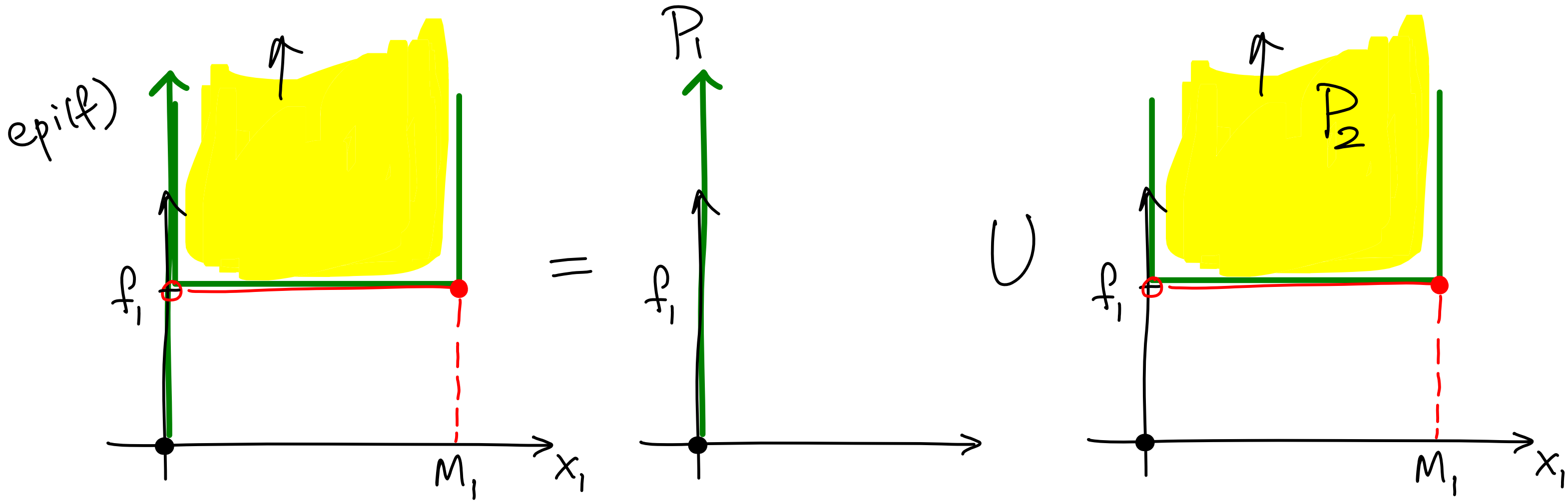



graph(f), i.e., $f(x_1)$ as drawn, is not b-MIP-r.

graph(f) is the union of origin (•) and and the half-open line segment (  ).

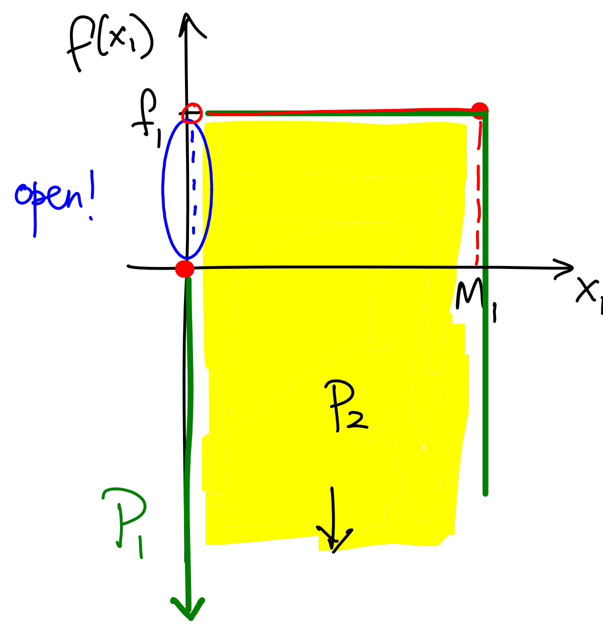
This second piece is not a polyhedron to start with.

But epi(f) turns out to be b-MIP-r here, and hypo(f) is not b-MIP-r at the same time.




Here $rec(P_1) = rec(P_2) = P_1$. Hence epi(f) is b-MIP-r.

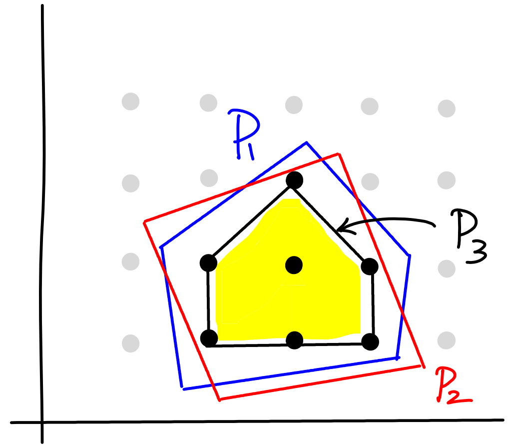
Here, hypo(f) is not b-MIP-r, as it is the union of $P_1$ and $P_2$, where $P_2$ is not a polyhedron (same reason as that for graph(f)).




What if $f_1 < 0$ here? Then hypo(f) will be b-MIP-r and epi(f) will not be b-MIP-r (the roles are reversed).

# Comparing Formulations

Q. What is a good/bad formulation?




We could say here $P_3$ is better than both $P_1$ and $P_2$, but cannot compare $P_1$ to $P_2$. More generally, we want to compare formulations with different sets (and hence different numbers) of extra variables. To this end, we need to introduce some basic results.

pronounced "Farkash"; feasibility of alternative systems

## Farkas' Lemma

① $\exists \bar{x} : A\bar{x} \le \bar{b}$ then

$A\bar{x} \le \bar{b}$ implies $\bar{a}^T \bar{x} \le \beta \iff$

$\exists \bar{u} \ge \bar{0}$ such that $\bar{u}^T A = \bar{a}^T$, $\bar{u}^T \bar{b} \le \beta$

multipliers using which we could derive $\bar{a}^T \bar{x} \le \beta$ from $A\bar{x} \le \bar{b}$

② $\exists \bar{x} : A\bar{x} \le \bar{b} \iff \nexists \bar{u} \ge \bar{0}, \ \bar{u}^T A = \bar{0}^T, \ \bar{u}^T \bar{b} < 0.$

(cannot derive $\bar{0}^T \bar{x} \le -1$ from $A\bar{x} \le \bar{b}$)