Linear arithmetic theories:
geometric procedures

Christoph Haase **Alessio Mansutti**

Today's lecture: geometric procedures

Quantifier elimination (Wednesday)**:**

$$\exists x : \Phi_{\mathsf{qf}}(x, \boldsymbol{y}) \;\equiv\; \Psi_{\mathsf{qf}}(\boldsymbol{y}) \qquad \text{qf: quantifier-free}$$

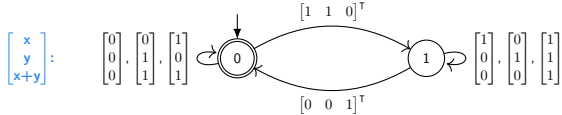
Today's lecture: geometric procedures

Quantifier elimination (Wednesday)**:**

$$\exists x : \Phi_{\mathsf{qf}}(x, \boldsymbol{y}) \;\equiv\; \Psi_{\mathsf{qf}}(\boldsymbol{y}) \qquad \text{qf: quantifier-free}$$

Automata techniques (yesterday)**:**

$z = x + y \quad \rightsquigarrow$



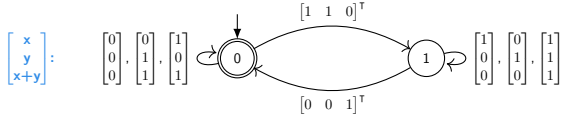
Today's lecture: geometric procedures

Quantifier elimination (Wednesday)**:**

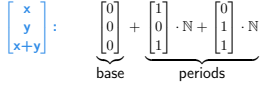
$$\exists x : \Phi_{\mathsf{qf}}(x, \boldsymbol{y}) \;\equiv\; \Psi_{\mathsf{qf}}(\boldsymbol{y}) \qquad \text{qf: quantifier-free}$$

Automata techniques (yesterday)**:**

$z = x + y \quad \rightsquigarrow$



Geometric procedures:

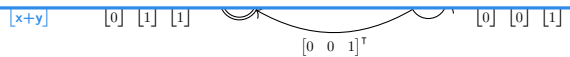


Today's lecture: geometric procedures

Quantifier elimination (Wednesday):

Content:

- Geometric procedure for Presburger arithmetic
- Presburger arithmetic extended with Kleene star
- The Vapnik–Chervonenkis (VC) dimension of linear arithmetics



Geometric procedures:

$$\begin{bmatrix} x \\ y \\ x+y \end{bmatrix} : \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{\text{base}} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \cdot \mathbb{N} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \cdot \mathbb{N}}_{\text{periods}}$$

What do we mean by geometric procedure?

In a broad sense, a **geometric procedure** takes an input formula Φ and

1. translates each atomic formula of Φ into a "geometric object" characterising all its solutions,
2. by applying operations of union, complementation and projection, derives a geometric object S_Φ characterizing the solutions of Φ.

What do we mean by geometric procedure?

In a broad sense, a **geometric procedure** takes an input formula Φ and

1. translates each atomic formula of Φ into a "geometric object" characterising all its solutions,
2. by applying operations of union, complementation and projection, derives a geometric object S_Φ characterizing the solutions of Φ.

"Geometric object" $=$ finite collection of vectors

What do we mean by geometric procedure?

In a broad sense, a **geometric procedure** takes an input formula Φ and

1. translates each atomic formula of Φ into a "geometric object" characterising all its solutions,
2. by applying operations of union, complementation and projection, derives a geometric object S_Φ characterizing the solutions of Φ.

"Geometric object" $=$ finite collection of vectors

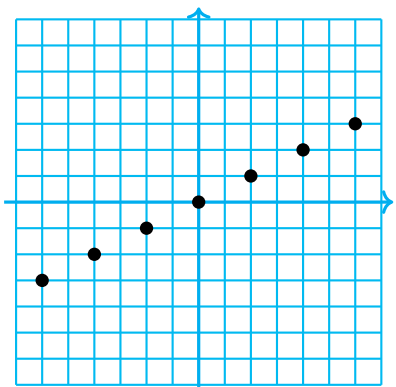
Computational features: Given S_Φ, computing a solution of Φ is in $\mathbf{PTIME}$.
If domain of the solutions is enumerable, then the solutions of Φ can be enumerated with polynomial delay.

Example in Presburger arithmetic

$$x \geq 4 \wedge \big(\big(\exists y : x = 2 \cdot y \big) \vee \big(\exists y : x - 1 = 3 \cdot y \big) \big)$$

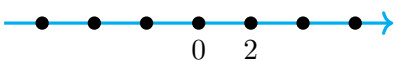
Example in Presburger arithmetic

$$x \geq 4 \wedge \big((\exists y : x = 2 \cdot y) \vee (\exists y : x - 1 = 3 \cdot y)\big)$$



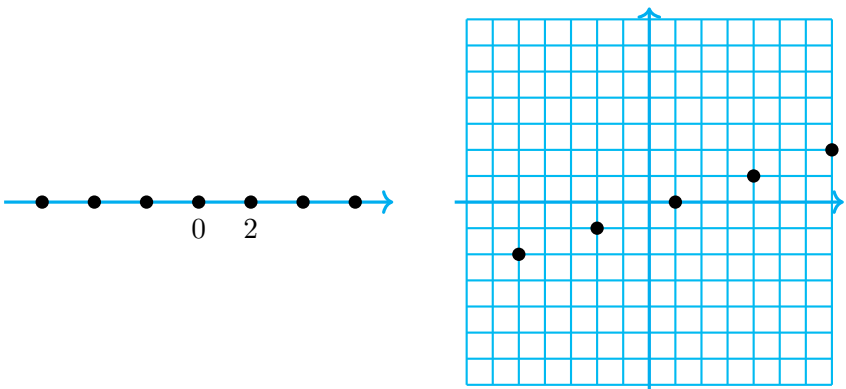
Example in Presburger arithmetic

$$x \geq 4 \wedge \big(\big(\exists y : x = 2 \cdot y \big) \vee \big(\exists y : x - 1 = 3 \cdot y \big) \big)$$



Example in Presburger arithmetic

$$x \geq 4 \wedge \big((\exists y : x = 2 \cdot y) \vee (\exists y : x - 1 = 3 \cdot y)\big)$$



Example in Presburger arithmetic

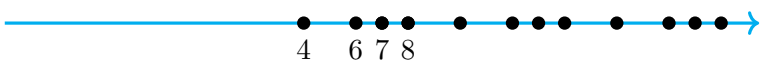
$$x \geq 4 \wedge \big(\big(\exists y : x = 2 \cdot y\big) \vee \big(\exists y : x - 1 = 3 \cdot y\big)\big)$$

Example in Presburger arithmetic

$$x \geq 4 \wedge \big((\exists y : x = 2 \cdot y) \vee (\exists y : x - 1 = 3 \cdot y)\big)$$

Example in Presburger arithmetic

$$x \geq 4 \wedge \big((\exists y : x = 2 \cdot y) \vee (\exists y : x - 1 = 3 \cdot y)\big)$$



$$\{4, 6, 7, 8\} + 6 \cdot \mathbb{N}$$

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.

Arithmetic progression

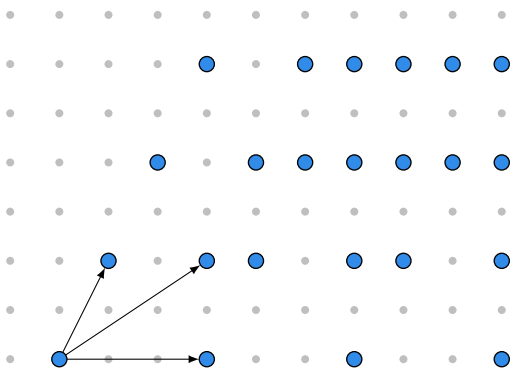
$b + p \cdot \mathbb{N}$, where $i \in \mathbb{N}$

b base point, p period

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



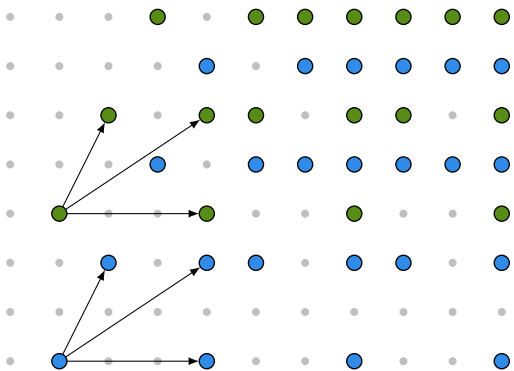
Linear set
(arithmetic progression
in multiple dimensions)

$L(\boldsymbol{b}, P)$, where $\boldsymbol{b}$ base
and $P = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\}$ periods

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



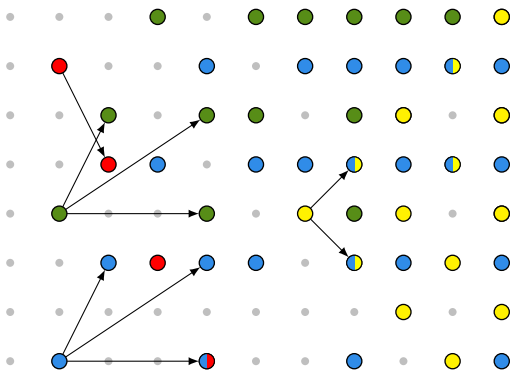
Hybrid linear set
(finite union of linear sets with same period)

$$L(B, P) \coloneqq \bigcup_{\boldsymbol{b} \in B} L(\boldsymbol{b}, P),$$

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



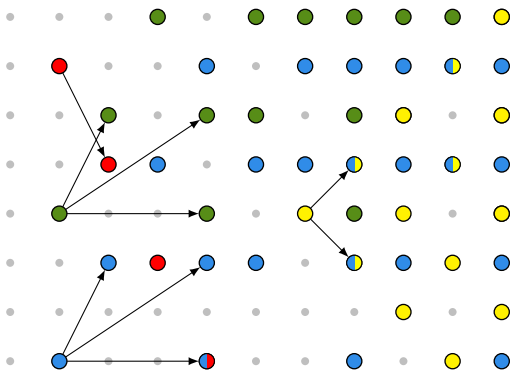
Semilinear set
(finite union of
hybrid linear sets)

$\bigcup_{i \in I} L(B_i, P_i)$,
where I finite set of indices

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



The set of solutions of a linear inequality over $\mathbb{Z}$ is semilinear.

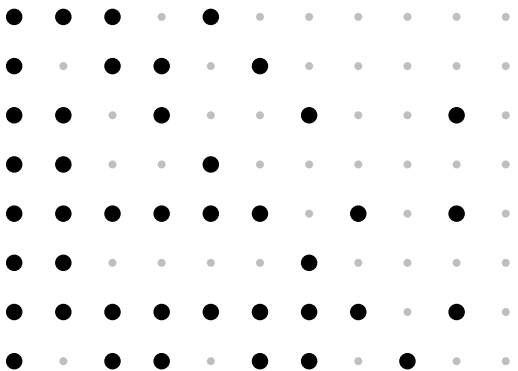
Semilinear sets are closed under

- union
- projection
- complementation

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



The set of solutions of a linear inequality over $\mathbb{Z}$ is semilinear.

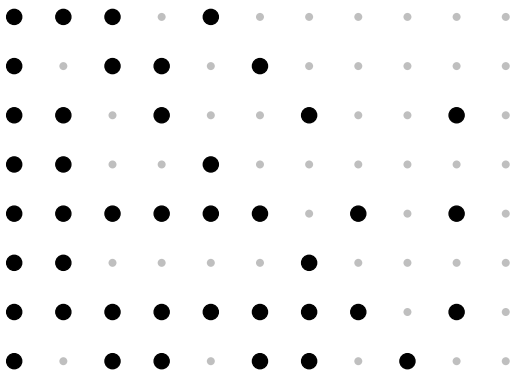
Semilinear sets are closed under

- union
- projection
- complementation

Semilinear sets

Theorem (Ginsburg & Spanier, 1966)

Sets definable in Linear integer arithmetic coincide with the family of semilinear sets.



The set of solutions of a linear inequality over $\mathbb{Z}$ is semilinear.

Semilinear sets are closed under

- union
- projection
- complementation

Further results are required:

- intersection of linear sets
- Carathéodory-type theorem

Solutions to linear inequalities over $\mathbb{Z}$ are semilinear

Theorem (von zur Gathen & Sieveking, 1978 — Lecture 2)

Consider $S \subseteq \mathbb{Z}^d$. Then, below (H) implies (V), but not vice versa:

(H) $S = \{\boldsymbol{x} \in \mathbb{Z}^d : A \cdot \boldsymbol{x} \geq \boldsymbol{c}\}$ *for some* $A \in \mathbb{Z}^{n \times d}$ *and* $\boldsymbol{c} \in \mathbb{Z}^m$

(V) $S = L(B, P)$ *for some finite sets* $B, P \subseteq \mathbb{Z}^d$.

Solutions to linear inequalities over $\mathbb{Z}$ are semilinear

Theorem (von zur Gathen & Sieveking, 1978 — Lecture 2)

Consider $S \subseteq \mathbb{Z}^d$. Then, below (H) implies (V), but not vice versa:

(H) $S = \{\boldsymbol{x} \in \mathbb{Z}^d : A \cdot \boldsymbol{x} \geq \boldsymbol{c}\}$ *for some* $A \in \mathbb{Z}^{n \times d}$ *and* $\boldsymbol{c} \in \mathbb{Z}^m$

(V) $S = L(B, P)$ *for some finite sets* $B, P \subseteq \mathbb{Z}^d$.

Proof for a single inequality $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$

1. find vectors $V = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{Z}^d$ such that (Minkowski–Weil theorem)
$$\operatorname{cone}(V) = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{a}^\intercal \cdot \boldsymbol{x} = 0\}$$

Solutions to linear inequalities over $\mathbb{Z}$ are semilinear

Theorem (von zur Gathen & Sieveking, 1978 — Lecture 2)

Consider $S \subseteq \mathbb{Z}^d$. Then, below (H) implies (V), but not vice versa:

(H) $S = \{\boldsymbol{x} \in \mathbb{Z}^d : A \cdot \boldsymbol{x} \geq \boldsymbol{c}\}$ *for some* $A \in \mathbb{Z}^{n \times d}$ *and* $\boldsymbol{c} \in \mathbb{Z}^m$

(V) $S = L(B, P)$ *for some finite sets* $B, P \subseteq \mathbb{Z}^d$.

Proof for a single inequality $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$

1. find vectors $V = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{Z}^d$ such that (Minkowski–Weil theorem)
$$\operatorname{cone}(V) = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{a}^\intercal \cdot \boldsymbol{x} = 0\}$$

2. find $\boldsymbol{x}^* \in \mathbb{Z}^d$ maximizing $-\boldsymbol{a}^\intercal \boldsymbol{x}$ subject to $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$ (ILP)

Solutions to linear inequalities over $\mathbb{Z}$ are semilinear

Theorem (von zur Gathen & Sieveking, 1978 — Lecture 2)

Consider $S \subseteq \mathbb{Z}^d$. Then, below (H) implies (V), but not vice versa:

(H) $S = \{\boldsymbol{x} \in \mathbb{Z}^d : A \cdot \boldsymbol{x} \geq \boldsymbol{c}\}$ *for some* $A \in \mathbb{Z}^{n \times d}$ *and* $\boldsymbol{c} \in \mathbb{Z}^m$

(V) $S = L(B, P)$ *for some finite sets* $B, P \subseteq \mathbb{Z}^d$.

Proof for a single inequality $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$

1. find vectors $V = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{Z}^d$ such that (Minkowski–Weil theorem)
   $$\operatorname{cone}(V) = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{a}^\intercal \cdot \boldsymbol{x} = 0\}$$
2. find $\boldsymbol{x}^* \in \mathbb{Z}^d$ maximizing $-\boldsymbol{a}^\intercal \boldsymbol{x}$ subject to $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$ (ILP)
3. $B \coloneqq \{\boldsymbol{x}^* + [V|\boldsymbol{a}] \cdot \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}^{k+1},\ \mathbf{0} \leq \boldsymbol{\lambda} < \mathbf{1}\} \cap \mathbb{Z}^d$ and $P \coloneqq \{\boldsymbol{a}, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$

Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)
3. eliminate from E and S entries corresponding to variables $\boldsymbol{y}$ (projection)
4. **return** $D \coloneqq \boldsymbol{c} + Q \cdot E$ and $R \coloneqq Q \cdot S$ $(A \cdot B \coloneqq \{A \cdot \boldsymbol{b} : \boldsymbol{b} \in B\})$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)
3. eliminate from E and S entries corresponding to variables $\boldsymbol{y}$ (projection)
4. **return** $D \coloneqq \boldsymbol{c} + Q \cdot E$ and $R \coloneqq Q \cdot S$ $(A \cdot B \coloneqq \{A \cdot \boldsymbol{b} : \boldsymbol{b} \in B\})$

$$L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q) \;=\; \boldsymbol{c} + \{Q \cdot \boldsymbol{x} : \boldsymbol{x} \in L(E, S)\}$$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)
3. eliminate from E and S entries corresponding to variables $\boldsymbol{y}$ (projection)
4. **return** $D \coloneqq \boldsymbol{c} + Q \cdot E$ and $R \coloneqq Q \cdot S$ $(A \cdot B \coloneqq \{A \cdot \boldsymbol{b} : \boldsymbol{b} \in B\})$

$$L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q) \;=\; \boldsymbol{c} + \{Q \cdot \boldsymbol{x} : \boldsymbol{x} \in L(E, S)\} \;=\; \boldsymbol{c} + Q \cdot L(E, S)$$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)
3. eliminate from E and S entries corresponding to variables $\boldsymbol{y}$ (projection)
4. **return** $D \coloneqq \boldsymbol{c} + Q \cdot E$ and $R \coloneqq Q \cdot S$ $(A \cdot B \coloneqq \{A \cdot \boldsymbol{b} : \boldsymbol{b} \in B\})$

$$\begin{aligned} L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q) &= \boldsymbol{c} + \{Q \cdot \boldsymbol{x} : \boldsymbol{x} \in L(E, S)\} = \boldsymbol{c} + Q \cdot L(E, S) \\ &= \boldsymbol{c} + Q \cdot (E + \{S \cdot \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{N}^n\}) \end{aligned}$$

Intersection of linear sets

Input: two pairs $(\boldsymbol{b}, P)$ and $(\boldsymbol{c}, Q)$ representing the linear sets $L(\boldsymbol{b}, P), L(\boldsymbol{c}, Q) \subseteq \mathbb{Z}^d$
Output: two finite sets $D, R \subseteq \mathbb{Z}^d$ such that $L(D, R) = L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q)$

1. Let $\mathfrak{S}$ be the system $\boldsymbol{b} + P \cdot \boldsymbol{x} = \boldsymbol{c} + Q \cdot \boldsymbol{y} \wedge \boldsymbol{x} \geq \mathbf{0} \wedge \boldsymbol{y} \geq \mathbf{0}$
2. compute $L(E, S)$ from $\mathfrak{S}$ (von zur Gathen & Sieveking)
3. eliminate from E and S entries corresponding to variables $\boldsymbol{y}$ (projection)
4. **return** $D \coloneqq \boldsymbol{c} + Q \cdot E$ and $R \coloneqq Q \cdot S$ $(A \cdot B \coloneqq \{A \cdot \boldsymbol{b} : \boldsymbol{b} \in B\})$

$$\begin{aligned} L(\boldsymbol{b}, P) \cap L(\boldsymbol{c}, Q) &= \boldsymbol{c} + \{Q \cdot \boldsymbol{x} : \boldsymbol{x} \in L(E, S)\} = \boldsymbol{c} + Q \cdot L(E, S) \\ &= \boldsymbol{c} + Q \cdot (E + \{S \cdot \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{N}^n\}) = L(\boldsymbol{c} + Q \cdot E, Q \cdot S) \end{aligned}$$

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i \in I}$ representing the semilinear set $S \coloneqq \bigcup_{i \in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m \in M}$ such that $\bigcup_{m \in M} L(E_m, S_m) = \overline{S}$

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S \coloneqq \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S \coloneqq \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

$$= \overline{\bigcup_{i\in I} \bigcup_{j\in J_i} L(\boldsymbol{c}_j, Q_j)}$$

where each Q_j is a set of linearly independent vectors

Carathéodory-type theorem

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S := \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

$$= \overline{\bigcup_{i\in I} \bigcup_{j\in J_i} L(\boldsymbol{c}_j, Q_j)}$$

where each Q_j is a set of linearly independent vectors

Carathéodory-type theorem

$$= \bigcap_{i\in I} \bigcap_{j\in J_i} \overline{L(\boldsymbol{c}_j, Q_j)}$$

De Morgan's law $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S := \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

$$= \overline{\bigcup_{i\in I} \bigcup_{j\in J_i} L(\boldsymbol{c}_j, Q_j)}$$

where each Q_j is a set of linearly independent vectors

Carathéodory-type theorem

$$= \bigcap_{i\in I} \bigcap_{j\in J_i} \overline{L(\boldsymbol{c}_j, Q_j)}$$

De Morgan's law $\overline{A \cup B} = \overline{A} \cap \overline{B}$

$$= \bigcap_{i\in I} \bigcap_{j\in J_i} \bigcup_{k\in K_j} L(D_k, R_k)$$

Complementation of linear sets with independent periods

Complementation of semilinear sets: overview

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S := \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$
Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

$$= \overline{\bigcup_{i\in I} \bigcup_{j\in J_i} L(\boldsymbol{c}_j, Q_j)}$$

where each Q_j is a set of linearly independent vectors

Carathéodory-type theorem

$$= \bigcap_{i\in I} \bigcap_{j\in J_i} \overline{L(\boldsymbol{c}_j, Q_j)}$$

De Morgan's law $\overline{A \cup B} = \overline{A} \cap \overline{B}$

$$= \bigcap_{i\in I} \bigcap_{j\in J_i} \bigcup_{k\in K_j} L(D_k, R_k)$$

Complementation of linear sets with independent periods

$$= \bigcup_{m\in M} L(E_m, S_m)$$

distribute $\cap$ over $\cup$, and intersection of linear sets

Making the period sets linearly independent

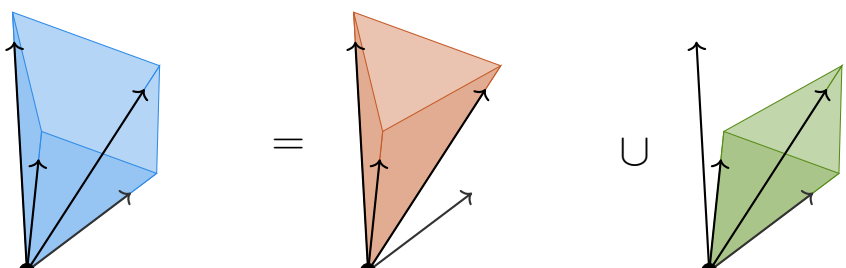
Theorem (Caratheodory, 1907)

- *Every point in conv*$(V) \subseteq \mathbb{R}^d$ *is a convex combination of* $d+1$ *vertices in* V*.*
- *Every point in cone*$(V) \subseteq \mathbb{R}^d$ *is a conic combination of* d *vertices in* V*.*

Making the period sets linearly independent

Theorem (Caratheodory, 1907)

- *Every point in conv*$(V) \subseteq \mathbb{R}^d$ *is a convex combination of* $d+1$ *vertices in* V.
- *Every point in cone*$(V) \subseteq \mathbb{R}^d$ *is a conic combination of* d *vertices in* V.



Making the period sets linearly independent

Theorem (Caratheodory, 1907)

- *Every point in conv$(V) \subseteq \mathbb{R}^d$ is a convex combination of $d+1$ vertices in V.*
- *Every point in cone$(V) \subseteq \mathbb{R}^d$ is a conic combination of d vertices in V.*

Theorem (Chistikov & Haase, 2016)

Consider $S = L(\boldsymbol{b}, P)$. Then, $S = \bigcup_{j \in J} L(C_j, Q_j)$ where vectors in $Q_j \subseteq P$ are linearly independent. The family $\{(C_j, Q_j)\}_{j \in J}$ can be effectively computed.

Making the period sets linearly independent

Theorem (Caratheodory, 1907)

- *Every point in conv$(V) \subseteq \mathbb{R}^d$ is a convex combination of $d+1$ vertices in V.*
- *Every point in cone$(V) \subseteq \mathbb{R}^d$ is a conic combination of d vertices in V.*

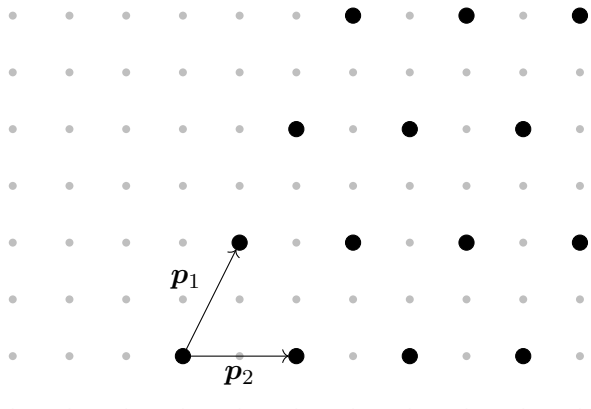
Theorem (Chistikov & Haase, 2016)

Consider $S = L(\boldsymbol{b}, P)$. Then, $S = \bigcup_{j \in J} L(C_j, Q_j)$ where vectors in $Q_j \subseteq P$ are linearly independent. The family $\{(C_j, Q_j)\}_{j \in J}$ can be effectively computed.

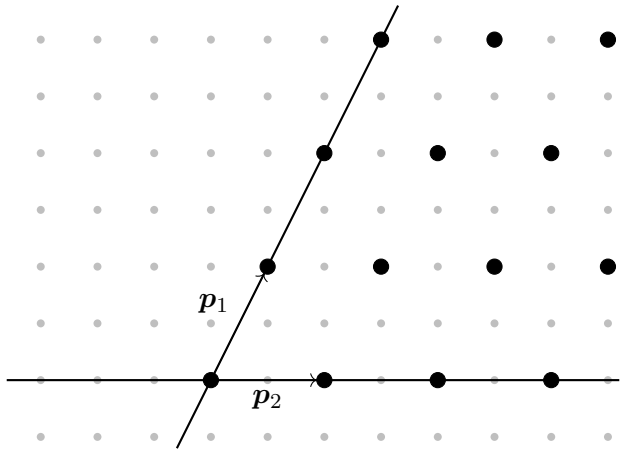
Proof.

Use Charathéodory's theorem on $\mathrm{cone}(P) \subseteq \mathbb{R}^d$ and discretise back. ∎

Complementing a linear set with linearly independent periods

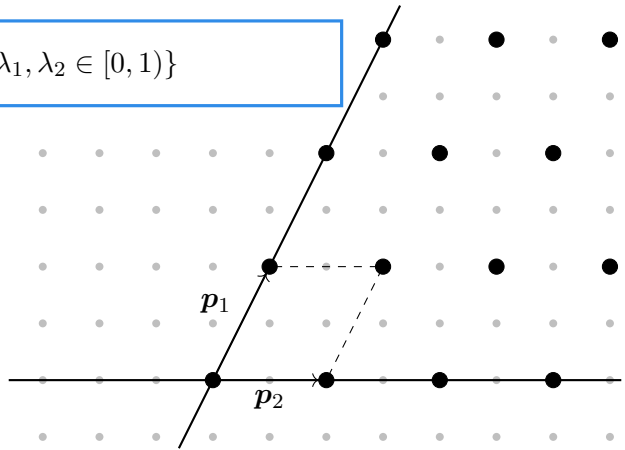


Complementing a linear set with linearly independent periods



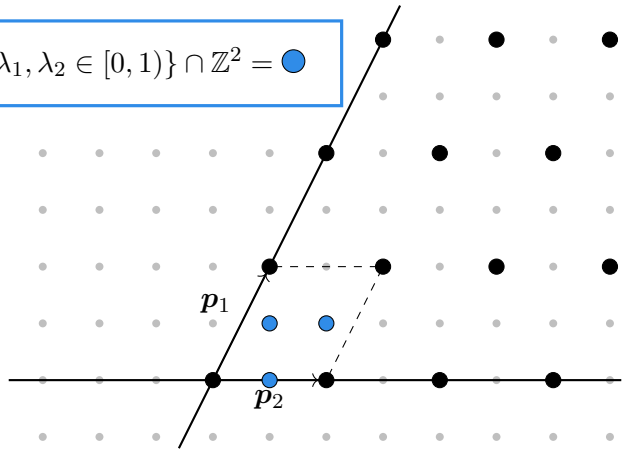
Complementing a linear set with linearly independent periods

$$\{\lambda_1 \cdot \boldsymbol{p}_1 + \lambda_2 \cdot \boldsymbol{p}_2 : \lambda_1, \lambda_2 \in [0, 1)\}$$



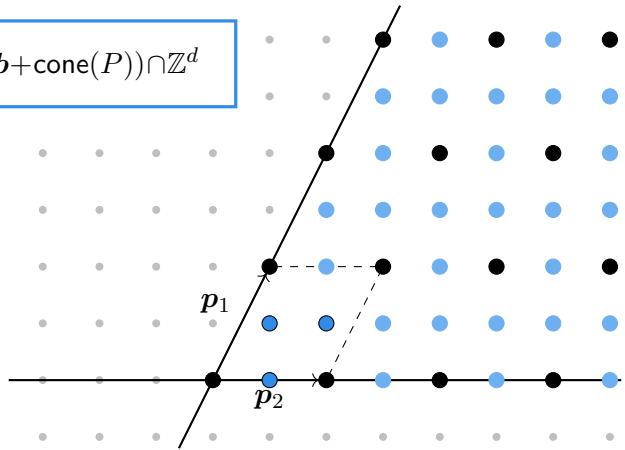
Complementing a linear set with linearly independent periods

$\{\lambda_1 \cdot \boldsymbol{p}_1 + \lambda_2 \cdot \boldsymbol{p}_2 : \lambda_1, \lambda_2 \in [0,1)\} \cap \mathbb{Z}^2 =$ ●



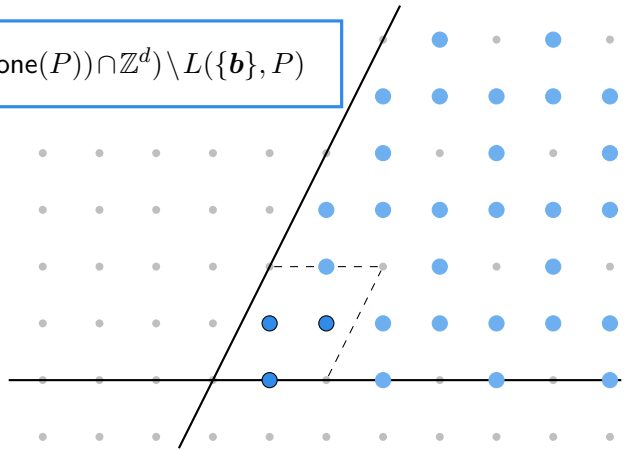
Complementing a linear set with linearly independent periods

$L(\{\boldsymbol{b}\} \cup \bullet, P) = (\boldsymbol{b} + \mathsf{cone}(P)) \cap \mathbb{Z}^d$

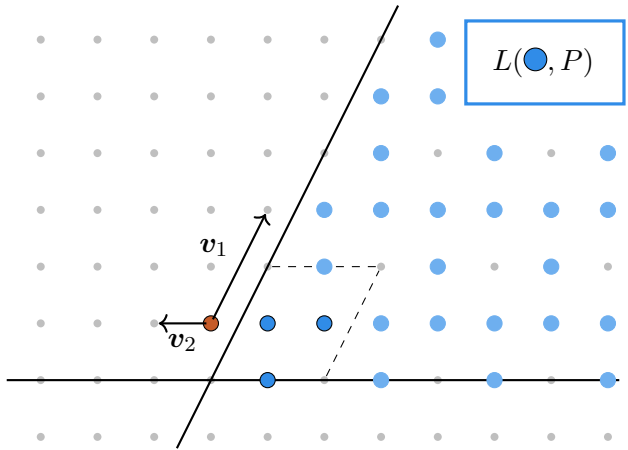


Complementing a linear set with linearly independent periods

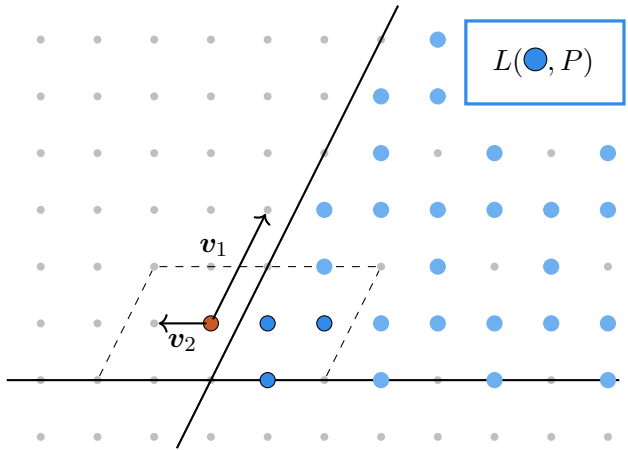
$$L(\bullet, P) = ((\boldsymbol{b} + \mathsf{cone}(P)) \cap \mathbb{Z}^d) \setminus L(\{\boldsymbol{b}\}, P)$$



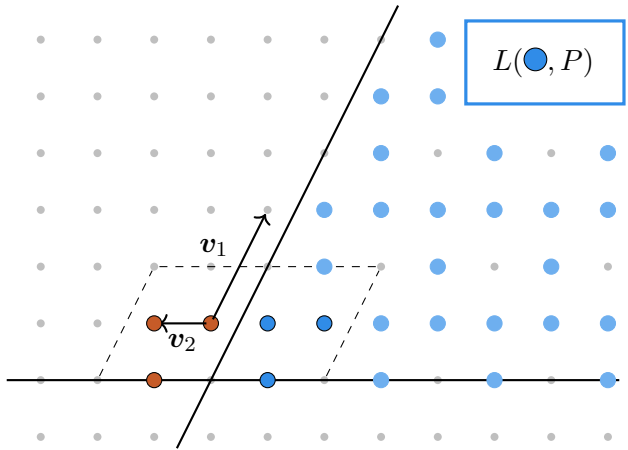
Complementing a linear set with linearly independent periods



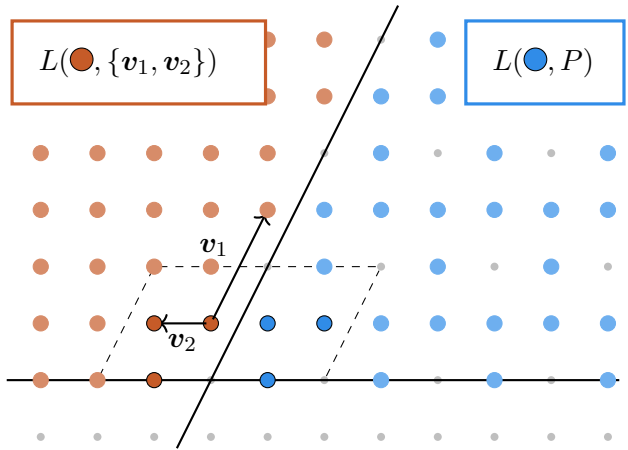
Complementing a linear set with linearly independent periods



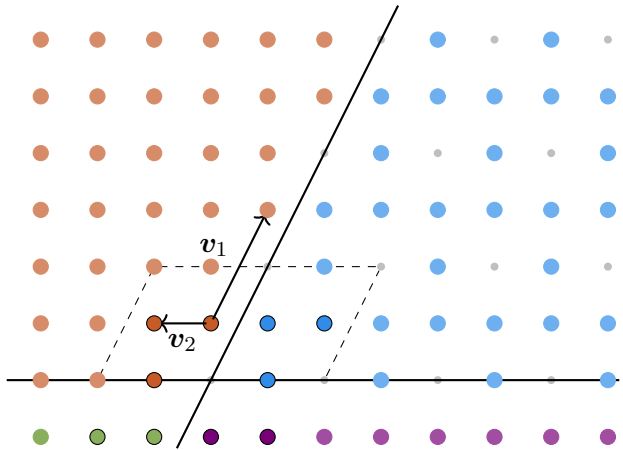
Complementing a linear set with linearly independent periods



Complementing a linear set with linearly independent periods



Complementing a linear set with linearly independent periods



Complementing a linear set with linearly independent periods (pseudocode)

Input: a pair $(\boldsymbol{b}, P)$ representing $L(\boldsymbol{b}, P)$; $P \subseteq \mathbb{Z}^d$ set of n linearly independent vectors

Output: a finite family $\{(C_i, Q_i)\}_{i \in I}$ such that $\bigcup_{i \in I} L(C_i, Q_i) = \overline{L(\boldsymbol{b}, P)}$

Complementing a linear set with linearly independent periods (pseudocode)

Input: a pair $(\boldsymbol{b}, P)$ representing $L(\boldsymbol{b}, P)$; $P \subseteq \mathbb{Z}^d$ set of n linearly independent vectors
Output: a finite family $\{(C_i, Q_i)\}_{i \in I}$ such that $\bigcup_{i \in I} L(C_i, Q_i) = \overline{L(\boldsymbol{b}, P)}$

1. $B \coloneqq \{\boldsymbol{b} + P \cdot \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \mathbb{R}^n,\ \mathbf{0} \leq \boldsymbol{\lambda} < \mathbf{1}\} \cap \mathbb{Z}^d$ (finite set)

2. **output** (C, P) where $C \coloneqq B \setminus L(\boldsymbol{b}, P)$ (membership queries)

3. compute a system $A \cdot \boldsymbol{x} \geq \boldsymbol{c}$ characterizing $\boldsymbol{b} + \mathrm{cone}(P)$ (Minkowski–Weyl)

4. **for every** $\mathfrak{S}$ obtained from $A \cdot \boldsymbol{x} \geq c$ by replacing some rows $\boldsymbol{a}^\intercal \cdot \boldsymbol{x} \geq c$ with their negations $\boldsymbol{a}^\intercal \boldsymbol{x} < c$ **do**

5. **output** (E, S) where $L(E, S)$ characterizes $\mathfrak{S}$ (von zur Gathen & Sieveking)

Complementation of semilinear sets: overview (again)

Input: a finite family $\{(\boldsymbol{b}_i, P_i)\}_{i\in I}$ representing the semilinear set $S \coloneqq \bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)$

Output: a finite family $\{(E_m, S_m)\}_{m\in M}$ such that $\bigcup_{m\in M} L(E_m, S_m) = \overline{S}$

$$\overline{\bigcup_{i\in I} L(\boldsymbol{b}_i, P_i)}$$

$$= \overline{\bigcup_{i\in I}\bigcup_{j\in J_i} L(\boldsymbol{c}_j, Q_j)}$$ where each Q_j is a set of linearly independent vectors

Carathéodory-type theorem

$$= \bigcap_{i\in I}\bigcap_{j\in J_i} \overline{L(\boldsymbol{c}_j, Q_j)}$$ De Morgan's law $\overline{A \cup B} = \overline{A} \cap \overline{B}$

$$= \bigcap_{i\in I}\bigcap_{j\in J_i}\bigcup_{k\in K_j} L(D_k, R_k)$$ Complementation of linear sets with independent periods

$$= \bigcup_{m\in M} L(E_m, S_m)$$ distribute $\cap$ over $\cup$, and intersection of linear sets

Complexity of geometric procedures

The complementation algorithm given in this lecture is highly suboptimal:
complementation followed by projection $=$ one **exponential** blow-up in the bit length

$$n \textbf{ complementations and projections} \implies 2^{2^{2^{\cdots^{2}}}} \Big\} O(n)$$

Complexity of geometric procedures

The complementation algorithm given in this lecture is highly suboptimal:

complementation followed by projection $=$ one **exponential** blow-up in the bit length

Theorem (Chistikov, Haase & Mansutti, 2022)

There is a $\mathbf{3EXPTIME}$ *geometric procedure for Presburger arithmetic.*

Complexity of geometric procedures

The complementation algorithm given in this lecture is highly suboptimal:
complementation followed by projection $=$ one **exponential** blow-up in the bit length

Theorem (Chistikov, Haase & Mansutti, 2022)

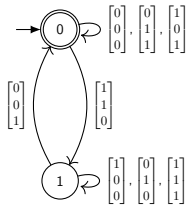
There is a $\mathbf{3EXPTIME}$ *geometric procedure for Presburger arithmetic.*

Quantifier elimination

$$\exists x : \Phi(x, \boldsymbol{y}) \;\rightsquigarrow\; \Psi(\boldsymbol{y})$$

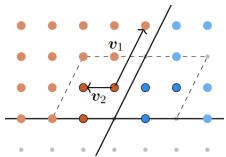
3EXPTIME

Automata



3EXPTIME

Geometry



3EXPTIME

Presburger arithmetic with Kleene stars

In regular languages, the Kleene star closes a language under repeated concatenations its words.

Presburger arithmetic with Kleene stars

In regular languages, the Kleene star closes a language under repeated concatenations its words. Any monoid $(M, e, \oplus)$ admit a similar notion: given $S \subseteq M$

$$S^0 = \{e\}$$

$$S^{n+1} = \{v \oplus w : v \in S^n \text{ and } w \oplus S\}$$

$$S^* = \bigcup_{k=0}^{\infty} S^k$$

Presburger arithmetic with Kleene stars

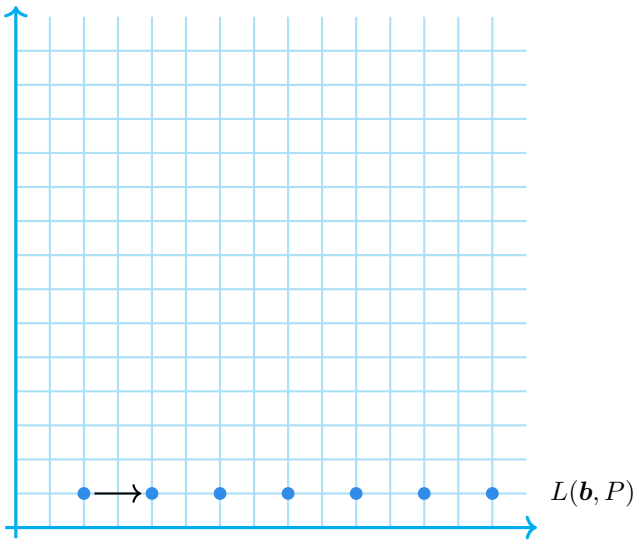
In regular languages, the Kleene star closes a language under repeated concatenations its words. Any monoid $(M, e, \oplus)$ admit a similar notion: given $S \subseteq M$

$$S^0 = \{e\}$$
$$S^{n+1} = \{v \oplus w : v \in S^n \text{ and } w \oplus S\}$$
$$S^* = \bigcup_{k=0}^{\infty} S^k$$

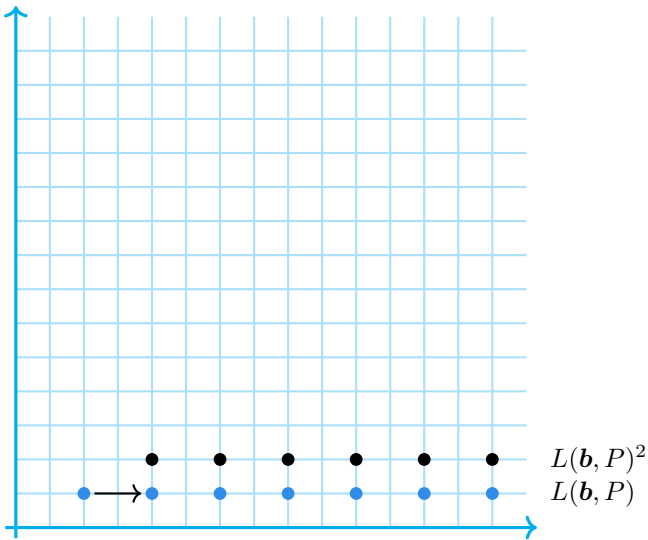
$(\mathbb{Z}^d, \mathbf{0}, +)$ is a monoid, so we can add the Kleene star to Presburger arithmetic: given a formula Φ representing the set $S \subseteq \mathbb{Z}^d$, the formula Φ^* represents the set

$$S^* = \bigcup_{k=0}^{\infty} \left\{ \sum_{i=0}^{k} \boldsymbol{v}_i : \boldsymbol{v}_0, \ldots, \boldsymbol{v}_i \in S \right\}$$

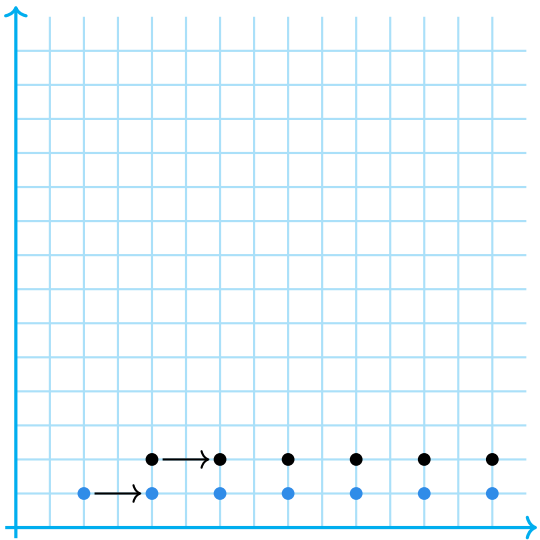
Elimination of Kleene stars, geometrically



Elimination of Kleene stars, geometrically



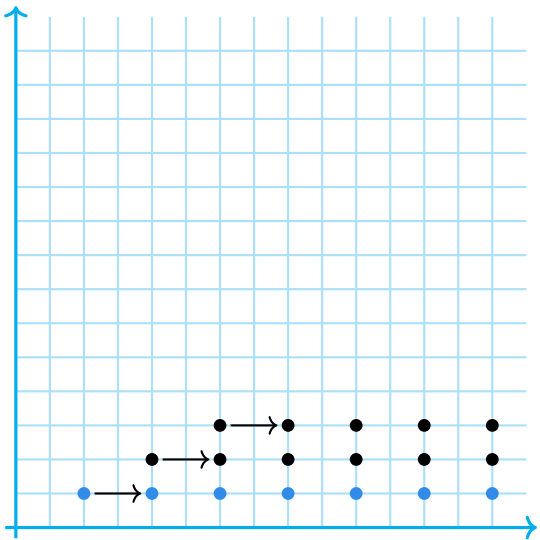
Elimination of Kleene stars, geometrically



$L(\boldsymbol{b}, P)^2 = L(2 \cdot \boldsymbol{b}, P)$

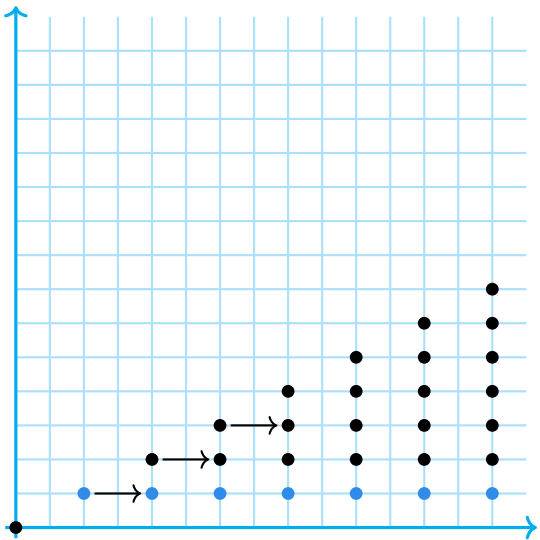
$L(\boldsymbol{b}, P)$

Elimination of Kleene stars, geometrically



$L(\boldsymbol{b}, P)^3 = L(3 \cdot \boldsymbol{b}, P)$
$L(\boldsymbol{b}, P)^2 = L(2 \cdot \boldsymbol{b}, P)$
$L(\boldsymbol{b}, P)$

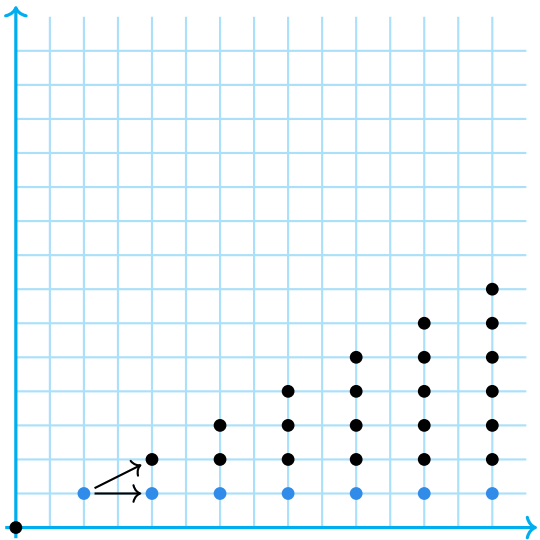
Elimination of Kleene stars, geometrically



$$L(\boldsymbol{b}, P)^* = \bigcup_{k=0}^{\infty} L(k \cdot \boldsymbol{b}, P)$$

$L(\boldsymbol{b}, P)$

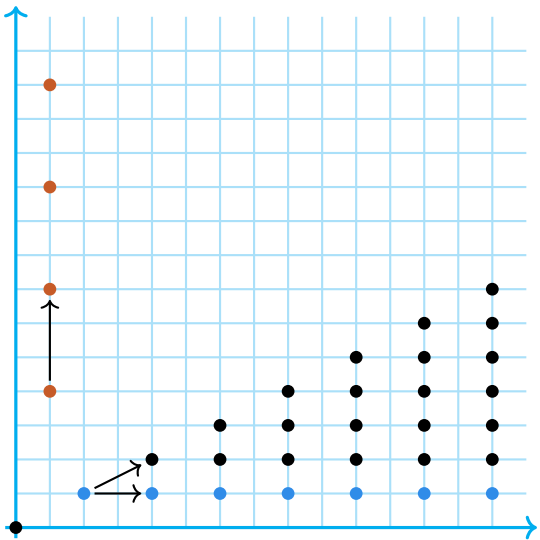
Elimination of Kleene stars, geometrically



$$L(\boldsymbol{b}, P)^* = \bigcup_{k=0}^{\infty} L(k \cdot \boldsymbol{b}, P)$$
$$= \{\mathbf{0}\} \cup L(\boldsymbol{b}, P \cup \{\boldsymbol{b}\})$$

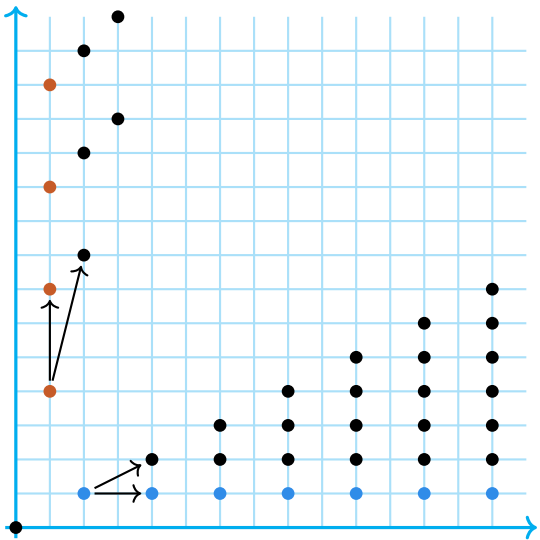
$L(\boldsymbol{b}, P)$

Elimination of Kleene stars, geometrically



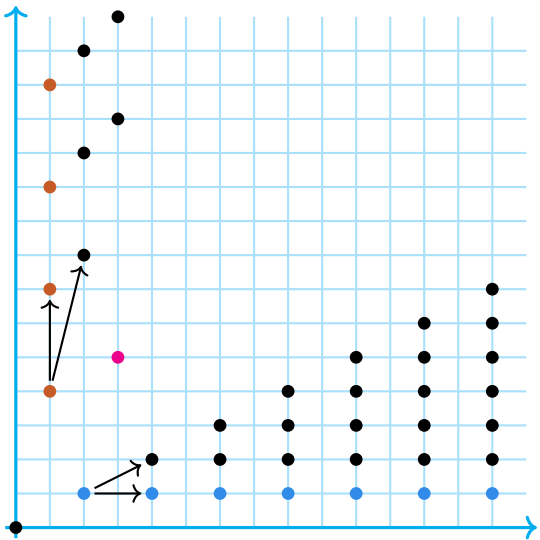
$$\Big(L(\boldsymbol{b}, P) \cup L(\boldsymbol{c}, Q)\Big)^*$$

Elimination of Kleene stars, geometrically



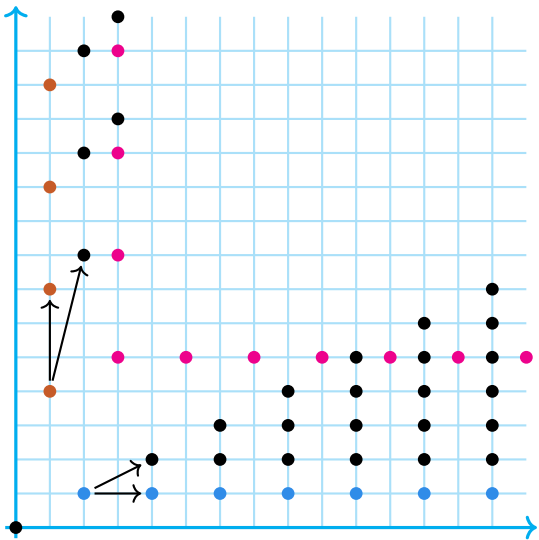
$$\Big(L(\boldsymbol{b}, P) \cup L(\boldsymbol{c}, Q)\Big)^*$$
$$= \{\mathbf{0}\} \cup L\left(\boldsymbol{b}, P \cup \{\boldsymbol{b}\}\right) \cup L\left(\boldsymbol{c}, Q \cup \{\boldsymbol{c}\}\right)$$

Elimination of Kleene stars, geometrically



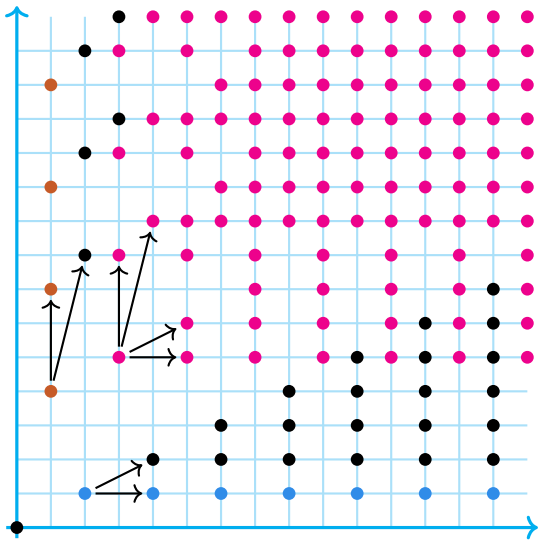
$$\Big(L(\boldsymbol{b}, P) \cup L(\boldsymbol{c}, Q)\Big)^*$$
$$= \{\mathbf{0}\} \cup L\left(\boldsymbol{b}, P \cup \{\boldsymbol{b}\}\right) \cup L\left(\boldsymbol{c}, Q \cup \{\boldsymbol{c}\}\right)$$

Elimination of Kleene stars, geometrically



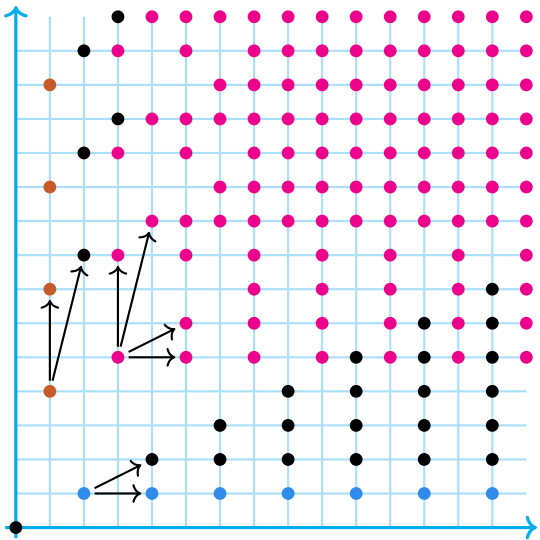
$$\Big(L(\boldsymbol{b}, P) \cup L(\boldsymbol{c}, Q)\Big)^*$$
$$= \{\mathbf{0}\} \cup L\left(\boldsymbol{b}, P \cup \{\boldsymbol{b}\}\right) \cup L\left(\boldsymbol{c}, Q \cup \{\boldsymbol{c}\}\right)$$

Elimination of Kleene stars, geometrically



$$\Big(L(\boldsymbol{b}, P) \cup L(\boldsymbol{c}, Q)\Big)^*$$
$$= \{\mathbf{0}\} \cup L\left(\boldsymbol{b}, P \cup \{\boldsymbol{b}\}\right) \cup L\left(\boldsymbol{c}, Q \cup \{\boldsymbol{c}\}\right)$$
$$\cup\ L\left(\boldsymbol{b} + \boldsymbol{c}, P \cup Q \cup \{\boldsymbol{b}, \boldsymbol{c}\}\right)$$

Elimination of Kleene stars, geometrically



$$\Big(L(\boldsymbol{b},P)\cup L(\boldsymbol{c},Q)\Big)^*$$
$$= \{\mathbf{0}\}\cup L\left(\boldsymbol{b},P\cup\{\boldsymbol{b}\}\right)\ \cup\ L\left(\boldsymbol{c},Q\cup\{\boldsymbol{c}\}\right)$$
$$\cup\ L\left(\boldsymbol{b}+\boldsymbol{c},P\cup Q\cup\{\boldsymbol{b},\boldsymbol{c}\}\right)$$

$$\left(\bigcup_{i\in I} L(\boldsymbol{b}_i,P_i)\right)^*$$
$$=$$
$$\bigcup_{J\subseteq I} L\left(\sum_{j\in J}\boldsymbol{b}_j,\ \bigcup_{j\in J}\Big(P_j\cup\{\boldsymbol{b}_j\}\Big)\right)$$

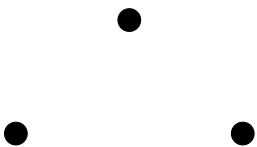
The VC dimension of linear arithmetics

Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.

Vapnik–Chervonenkis dimension

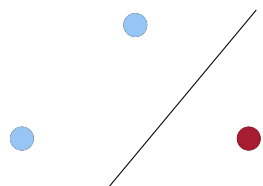
Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

Vapnik–Chervonenkis dimension

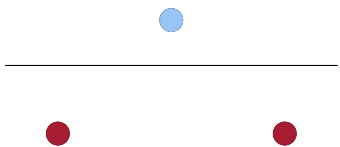
Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

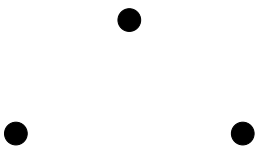
Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.

In how many ways we can classify these three points using one straight line?

Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.

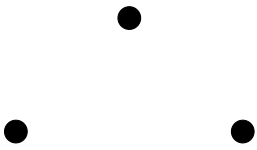


In how many ways we can classify these three points using one straight line?

2^3

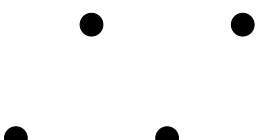
Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

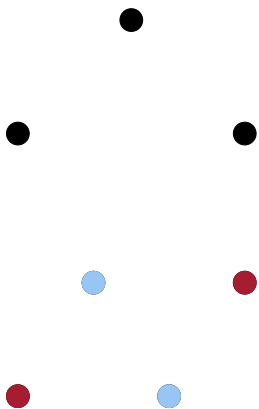
$$2^3$$



In how many ways we can classify four abritrary points using one straight line?

Vapnik–Chervonenkis dimension

Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

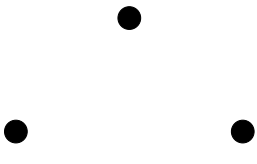
$$2^3$$

In how many ways we can classify four abritrary points using one straight line?

$$< 2^4$$

Vapnik–Chervonenkis dimension

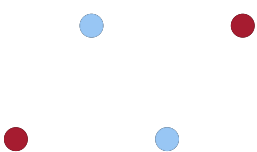
Measure of the capacity ($\sim$ expressiveness) of a set of functions that can be learned by a classification model.



In how many ways we can classify these three points using one straight line?

2^3

The VC dimension of a straight line classifier is 3.



In how many ways we can classify four abritrary points using one straight line?

$< 2^4$

Vapnik–Chervonenkis dimension: formal definition

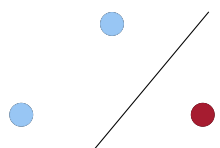
Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.

Vapnik–Chervonenkis dimension: formal definition

Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.



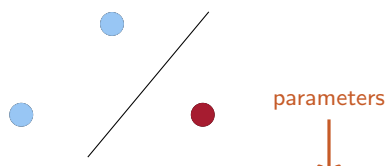
$$X = \mathbb{Z}^2, \quad \mathcal{F} = \{\{(x, y) : y \leq ax + b\} : a, b \in \mathbb{R}\}$$

The VC dimension of $\mathcal{F}$ is 3.

Vapnik–Chervonenkis dimension: formal definition

Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.



$$X = \mathbb{Z}^2, \quad \mathcal{F} = \{\{(x, y) : y \leq ax + b\} : a, b \in \mathbb{R}\}$$

The VC dimension of $\mathcal{F}$ is 3.

Statistical learning theory and bounds from the VC dimension

Consider an input space X and the output space $\{-1, +1\}$.

Assume an unknown probability distributions ρ over the product $X \times \{-1, +1\}$.

Goal: given $t_1, \ldots, t_m$ (training data) independently sampled according to ρ
obtain a classifier $f \colon X \to \{-1, +1\}$
minimizing as much as possible $R(f) \coloneqq \mathbb{E}_\rho[\mathbf{1}_{f(X) \neq Y}]$.

Statistical learning theory and bounds from the VC dimension

Consider an input space X and the output space $\{-1, +1\}$.

Assume an unknown probability distributions ρ over the product $X \times \{-1, +1\}$.

Goal: given $t_1, \ldots, t_m$ (training data) independently sampled according to ρ
obtain a classifier $f \colon X \to \{-1, +1\}$
minimizing as much as possible $R(f) \coloneqq \mathbb{E}_\rho[\mathbf{1}_{f(X) \neq Y}]$.

Fix a family $\mathcal{F} \subseteq X \to \{-1, +1\}$. A (binary) learning algorithm $\mathcal{A}$ for $\mathcal{F}$ is a procedure that on input training data returns a classifier $f \in \mathcal{F}$.

Statistical learning theory and bounds from the VC dimension

Consider an input space X and the output space $\{-1, +1\}$.

Assume an unknown probability distributions ρ over the product $X \times \{-1, +1\}$.

Goal: given $t_1, \ldots, t_m$ (training data) independently sampled according to ρ
obtain a classifier $f \colon X \to \{-1, +1\}$
minimizing as much as possible $R(f) \coloneqq \mathbb{E}_\rho[\mathbf{1}_{f(X) \neq Y}]$.

Fix a family $\mathcal{F} \subseteq X \to \{-1, +1\}$. A (binary) learning algorithm $\mathcal{A}$ for $\mathcal{F}$ is a procedure that on input training data returns a classifier $f \in \mathcal{F}$.

Question: let $\epsilon, \delta > 0$. How big the training data T needs to be to achieve

$$\Pr[R(\mathcal{A}(T)) - \inf_{h \in \mathcal{F}} R(h) \geq \epsilon] < \delta$$

assuming that $\mathcal{A}(T)$ probabilistically converges to the optimal $h \in \mathcal{F}$ as T increases.

Statistical learning theory and bounds from the VC dimension

Consider an input space X and the output space $\{-1, +1\}$

Assum }.

Theorem (Vapnik & Chervonenkis, 1971; Ehrenfeucht, Haussler, Kearns & Valiant, 1989)

Goal:

If $\mathcal{F}$ has VC dimension D, then $\Theta\left(\frac{D + \frac{1}{\ln \delta}}{\epsilon}\right)$ samples do the job.

Fix a family $\mathcal{F} \subseteq X \to \{-1, +1\}$. A (binary) learning algorithm $\mathcal{A}$ for $\mathcal{F}$ is a procedure that on input training data returns a classifier $f \in \mathcal{F}$.

Question: let $\epsilon, \delta > 0$. How big the training data T needs to be to achieve

$$\Pr[R(\mathcal{A}(T)) - \inf_{h \in \mathcal{F}} R(h) \geq \epsilon] < \delta$$

assuming that $\mathcal{A}(T)$ probabilistically converges to the optimal $h \in \mathcal{F}$ as T increases.

Vapnik–Chervonenkis dimension for FO formulae

Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.

Vapnik–Chervonenkis dimension for FO formulae

Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.

Let $\Phi(\boldsymbol{x}, \boldsymbol{y})$ be a FO formula from the theory of the structure $\mathcal{M}$, with universe M, where $\boldsymbol{x}$ are n free object variables and $\boldsymbol{y}$ are m free parameter variables.

$$\mathcal{F}_\Phi = \{\{\boldsymbol{x} \in M^n : \mathcal{M} \models \Phi(\boldsymbol{x}, \boldsymbol{y})\} : \boldsymbol{y} \in M^m\}.$$

The VC dimension of Φ is the VC dimension of $\mathcal{F}_\Phi$.

Vapnik–Chervonenkis dimension for FO formulae

Consider a set X and a family $\mathcal{F} \subseteq 2^X$.

- $\mathcal{F}$ shatters $A \subseteq X$ whenever $\{A \cap F : F \in \mathcal{F}\} = 2^A$.
- The VC dimension of $\mathcal{F}$ is the cardinality of the largest set A shattered by $\mathcal{F}$.

Let $\Phi(\boldsymbol{x}, \boldsymbol{y})$ be a FO formula from the theory of the structure $\mathcal{M}$, with universe M, where $\boldsymbol{x}$ are n free object variables and $\boldsymbol{y}$ are m free parameter variables.

$$\mathcal{F}_\Phi = \{\{\boldsymbol{x} \in M^n : \mathcal{M} \models \Phi(\boldsymbol{x}, \boldsymbol{y})\} : \boldsymbol{y} \in M^m\}.$$

The VC dimension of Φ is the VC dimension of $\mathcal{F}_\Phi$.

Theorem (see e.g. Bartlett, Maiorov, Meir, 1998)

The VC dimension of a neural network is polynomial in its size.

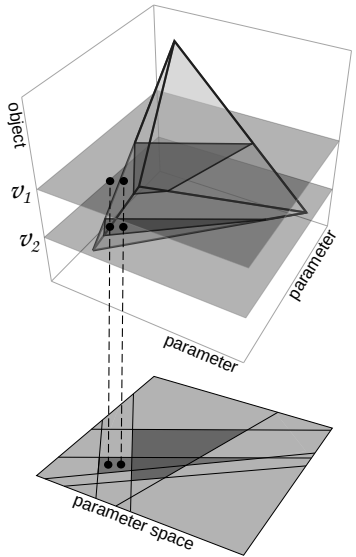
The VC dimension of linear arithmetic theories

Theorem (Chistikov, Haase & Mansutti, 2022)

Object	VC dimension
$\bigcup_{i\in I}\left(K(V_i,W_i)\setminus\bigcup_{j\in J_i}K(V_j,W_j)\right)\subseteq\mathbb{R}^d$	*polynomial*
$\bigcup_{i\in I}L(B_i,P_i)\subseteq\mathbb{Z}^d$	*only exponential in the dimension* d
Φ *in linear real arithmetic*	*exponential in* $\langle\Phi\rangle$
Φ *in Presburger arithmetic*	*doubly exponential in* $\langle\Phi\rangle$

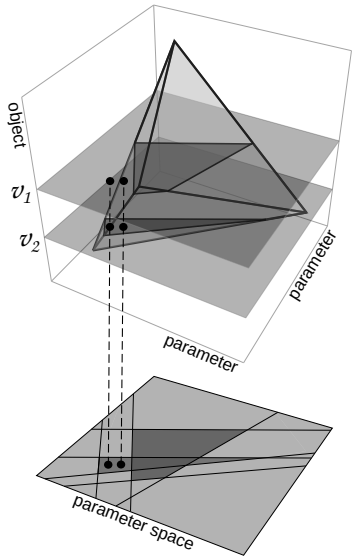
where $K(V,W) \coloneqq \mathrm{conv}(V)+\mathrm{cone}(W)$.

VC dimension upper bounds: counting regions in the parametric space



Consider a polyhedron $\Phi(\boldsymbol{x}, \boldsymbol{y}) \coloneqq A \cdot (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{b}$.
$\boldsymbol{x}$: object variables; $\boldsymbol{y}$: parameter variables

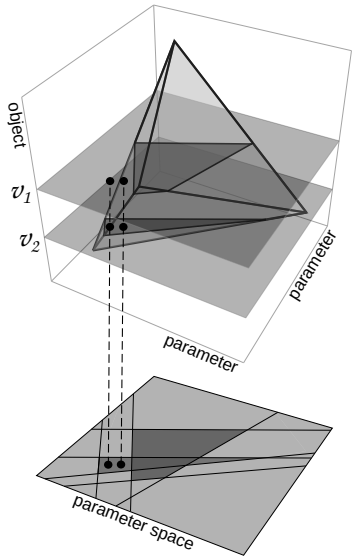
VC dimension upper bounds: counting regions in the parametric space



Consider a polyhedron $\Phi(\boldsymbol{x}, \boldsymbol{y}) \coloneqq A \cdot (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{b}$.
$\boldsymbol{x}$: object variables; $\boldsymbol{y}$: parameter variables

- given $\boldsymbol{v} \in \mathbb{R}^n$, $\Phi(\boldsymbol{v}, \boldsymbol{y})$ is a polyhedron that lives in the parameter space
- supporting hyperplanes characterising $\Phi(\boldsymbol{v}, \boldsymbol{y})$ split the parameter space into regions

VC dimension upper bounds: counting regions in the parametric space



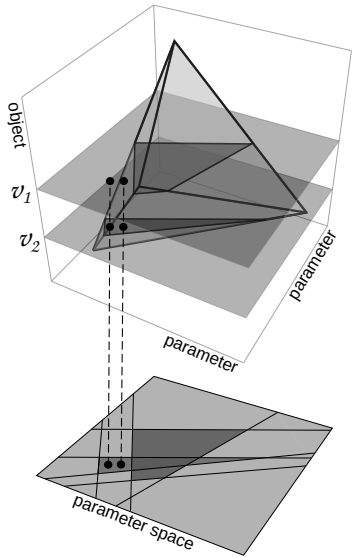
Consider a polyhedron $\Phi(\boldsymbol{x}, \boldsymbol{y}) \coloneqq A \cdot (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{b}$.
$\boldsymbol{x}$: object variables; $\boldsymbol{y}$: parameter variables

- given $\boldsymbol{v} \in \mathbb{R}^n$, $\Phi(\boldsymbol{v}, \boldsymbol{y})$ is a polyhedron that lives in the parameter space
- supporting hyperplanes characterising $\Phi(\boldsymbol{v}, \boldsymbol{y})$ split the parameter space into regions

Assume $\mathcal{F}_\Phi$ shatters a set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. Then,

$2^k \leq$ num. of regions induced by $\Phi(\boldsymbol{v}_1, \boldsymbol{y}), \ldots, \Phi(\boldsymbol{v}_k, \boldsymbol{y})$

VC dimension upper bounds: counting regions in the parametric space



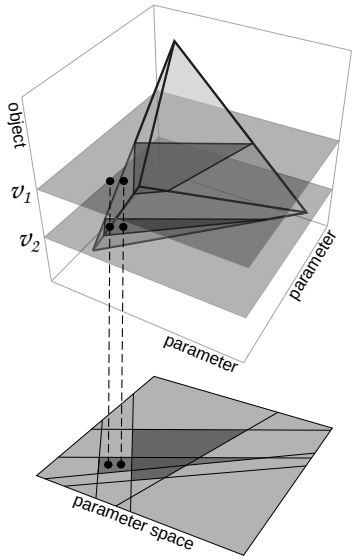
Consider a polyhedron $\Phi(\boldsymbol{x}, \boldsymbol{y}) \coloneqq A \cdot (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{b}$.
$\boldsymbol{x}$: object variables; $\boldsymbol{y}$: parameter variables

- given $\boldsymbol{v} \in \mathbb{R}^n$, $\Phi(\boldsymbol{v}, \boldsymbol{y})$ is a polyhedron that lives in the parameter space
- supporting hyperplanes characterising $\Phi(\boldsymbol{v}, \boldsymbol{y})$ split the parameter space into regions

Assume $\mathcal{F}_\Phi$ shatters a set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. Then,

$$2^k \leq \text{num. of regions induced by } \Phi(\boldsymbol{v}_1, \boldsymbol{y}), \ldots, \Phi(\boldsymbol{v}_k, \boldsymbol{y})$$
$$\leq k^{\mathrm{poly}\langle\Phi\rangle}$$

VC dimension upper bounds: counting regions in the parametric space



Consider a polyhedron $\Phi(\boldsymbol{x}, \boldsymbol{y}) \coloneqq A \cdot (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{b}$.
$\boldsymbol{x}$: object variables; $\boldsymbol{y}$: parameter variables

- given $\boldsymbol{v} \in \mathbb{R}^n$, $\Phi(\boldsymbol{v}, \boldsymbol{y})$ is a polyhedron that lives in the parameter space
- supporting hyperplanes characterising $\Phi(\boldsymbol{v}, \boldsymbol{y})$ split the parameter space into regions

Assume $\mathcal{F}_\Phi$ shatters a set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. Then,

$$2^k \leq \text{num. of regions induced by } \Phi(\boldsymbol{v}_1, \boldsymbol{y}), \ldots, \Phi(\boldsymbol{v}_k, \boldsymbol{y})$$
$$\leq k^{\mathrm{poly}\langle\Phi\rangle}$$

$\implies$ we can upper bound k with respect to $\langle\Phi\rangle$

What we have learned in this course

Convex geometry

- Convex sets, cones, polyhedra
- Faces of polyhedra
- Dimension of polyhedra
- Farkas' lemma
- Minkowski—Weyl theorem

Linear programming and Integer Linear Programming

- Simplex method
- Linear programming is in $\mathbf{PTIME}$
- Integer Linear programming is $\mathbf{NP}$-complete
- Branch and bound
- Randomized rounding

Algorithmic paradigms for arithmetic theories

- Quantifier elimination
- Automata-based procedures
- Geometric procedures
- Extensions of Presburger arithmetic: counting, exponentiation, Kleene star
- Learning theory: the VC dimension of linear arithmetic theories

A Survival Guide to Presburger Arithmetic

Christoph Haase, University of Oxford, UK

The first-order theory of the integers with addition and order, commonly known as Presburger arithmetic, has been a central topic in mathematical logic and computer science for almost 90 years. Presburger arithmetic has been the starting point for numerous lines of research in automata theory, model theory and discrete geometry. In formal verification, Presburger arithmetic is the first-choice logic to represent and reason about systems with infinitely many states. This article provides a broad yet concise overview over the history, decision procedures, extensions and geometric properties of Presburger arithmetic.

1. A VERY SHORT HISTORY OF PRESBURGER ARITHMETIC

Around the 1920s of the last millennium, David Hilbert together with his doctoral student Wilhelm Ackermann began to pursue what is nowadays known as *Hilbert's program*. The goal of this program was to create a formal system that would allow for providing solid foundations for all of mathematics. The means to achieve this goal was to use mathematical logic as an unambiguous language in which all mathematical statements could be formalised and manipulated according to a well-defined axiomatic system. In addition to asking for consistency and completeness, Hilbert also required that it should be possible to verify or falsify the truth of any given mathematical statement in a finite number of steps within this formal system. This requirement gave rise to the *Entscheidungsproblem (decision problem)* that was introduced by Hilbert and Ackermann in their book *Grundzüge der Theoretischen Logik (Principles of Mathematical Logic)* published in 1928, see [Hilbert and Ackermann 1950] for an English translation. The Entscheidungsproblem demands an algorithm that given a sentence in first-order logic together with a finite number of axioms allows for deciding whether that sentence is valid, i.e., holds in any structure satisfying the given axioms.

After studying the *Principles of Mathematical Logic* and related work, Alfred Tarski approached his student Mojżesz Presburger and asked him to investigate the completeness of a particular theory capturing a limited fragment of number theory. A couple of months later, Presburger showed in his Master's thesis the completeness

Fig. 1. Presburger's student card from the University of Warsaw, Poland.

A Survival Guide to Presburger Arithmetic

Christoph Haase, University of Oxford, UK

The first-order theory of the integers with addition and order, commonly known as Presburger arithmetic, has been a central topic in mathematical logic and computer science for almost 90 years. Presburger arithmetic has been the starting point for numerous lines of research in automata theory, model theory and discrete geometry. In formal verification, Presburger arithmetic is the first-choice logic to represent and reason about systems with infinitely many states. This article provides a broad yet concise overview over the history, decision procedures, extensions and geometric properties of Presburger arithmetic.

1. A VERY SHORT HISTORY OF PRESBURGER ARITHMETIC

Around the 1920s of the last millennium, David Hilbert together with his doctoral student Wilhelm Ackermann began to pursue what is nowadays known as *Hilbert's program*. The goal of this program was to create a formal system that would allow for providing solid foundations for all of mathematics. The means to achieve this goal was to use mathematical logic as an unambiguous language in which all mathematical statements could be formalised and manipulated according to a well-defined axiomatic system. In addition to asking for consistency and completeness, Hilbert also required that it should be possible to verify or falsify the truth of any given mathematical statement in a finite number of steps within this formal system. This requirement gave rise to the *Entscheidungsproblem (decision problem)* that was introduced by Hilbert and Ackermann in their book *Grundzüge der Theoretischen Logik (Principles of Mathematical Logic)* published in 1928, see [Hilbert and Ackermann 1950] for an English translation. The Entscheidungsproblem demands an algorithm that given a sentence in first-order logic together with a finite number of axioms allows for deciding whether that sentence is valid, i.e., holds in any structure satisfying the given axioms.

After studying the *Principles of Mathematical Logic* and related work, Alfred Tarski approached his student Mojżesz Presburger and asked him to investigate the completeness of a particular theory capturing a limited fragment of number theory. A couple of months later, Presburger showed in his Master's thesis the completeness

Fig. 1. Presburger's student card from the University of Warsaw, Poland.

Our groups are looking for new interns, PhD students and postdocs!

A Survival Guide to Presburger Arithmetic

Christoph Haase, University of Oxford, UK

The first-order theory of the integers with addition and order, commonly known as Presburger arithmetic, has been a central topic in mathematical logic and computer science for almost 90 years. Presburger arithmetic has been the starting point for numerous lines of research in automata theory, model theory and discrete geometry. In formal verification, Presburger arithmetic is the first-choice logic to represent and reason about systems with infinitely many states. This article provides a broad yet concise overview over the history, decision procedures, extensions and geometric properties of Presburger arithmetic.

1. A VERY SHORT HISTORY OF PRESBURGER ARITHMETIC

Around the 1920s of the last millennium, David Hilbert together with his doctoral student Wilhelm Ackermann began to pursue what is nowadays known as *Hilbert's program*. The goal of this program was to create a formal system that would allow for providing solid foundations for all of mathematics. The means to achieve this goal was to use mathematical logic as an unambiguous language in which all mathematical statements could be formalised and manipulated according to a well-defined axiomatic system. In addition to asking for consistency and completeness, Hilbert also required that it should be possible to verify or falsify the truth of any given mathematical statement in a finite number of steps within this formal system. This requirement gave rise to the *Entscheidungsproblem (decision problem)* that was introduced by Hilbert and Ackermann in their book *Grundzüge der Theoretischen Logik (Principles of Mathematical Logic)* published in 1928, see [Hilbert and Ackermann 1950] for an English translation. The Entscheidungsproblem demands an algorithm that given a sentence in first-order logic together with a finite number of axioms allows for deciding whether that sentence is valid, i.e., holds in any structure satisfying the given axioms.

After studying the *Principles of Mathematical Logic* and related work, Alfred Tarski approached his student Mojżesz Presburger and asked him to investigate the completeness of a particular theory capturing a limited fragment of number theory. A couple of months later, Presburger showed in his Master's thesis the completeness

Fig. 1. Presburger's student card from the University of Warsaw, Poland.

Our groups are looking for new interns, PhD students and postdocs!

Thank you!

christoph.haase@cs.ox.ac.uk
alessio.mansutti@imdea.org