Introduction to algebraic complexity theory and how geometry enters

Christian Ikenmeyer

Algebraic Complexity Theory Workshop at ICALP 2023

Agenda

Agenda

Algebraic algorithms

- Fast Fourier transform, fast matrix multiplication, ...
- Solving systems of linear equations
- Solving systems of polynomial equations: Gröbner bases
- Coding Theory: Reed-Muller codes, ...
- Number theory: Euclidean algorithm, Chinese Remainder Theorem, ...

Analyzing running time of algebraic algorithms:

- Number of arithmetic operations
- Size/growth/precision of the numbers

Arithmetization

Computation modulo 2: The field $\mathbb{F}_2$

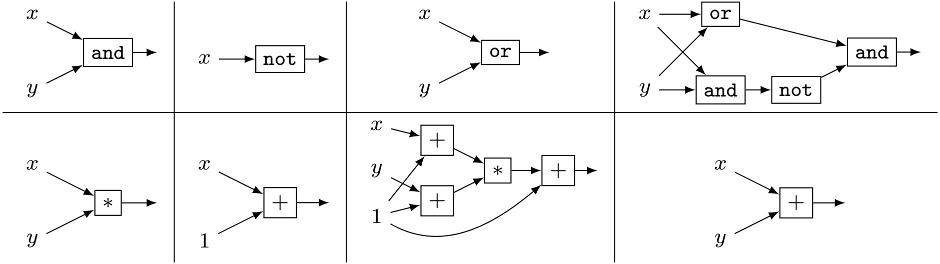
$*$	0	1
0	0	0
1	0	1

Boolean "and"

$+$	0	1
0	0	1
1	1	0

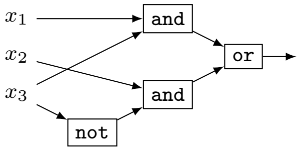
Boolean "xor"

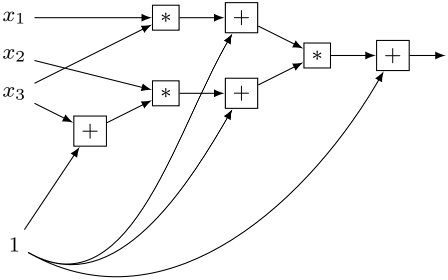
Translate Boolean circuit using $\{\texttt{and}, \texttt{or}, \texttt{not}\} \rightsquigarrow$ algebraic circuit using $\{+, *\}$:



During the translation the circuit only grows in size by at most a factor of 4.

Example





$(x_1$ **and** $x_3)$ **or** $(x_2$ **and** **not**$(x_3))$

$x_1 x_2 x_3^2 + x_1 x_2 x_3 + x_1 x_3 + x_2 x_3 + x_2$

Infinite fields

- Algebraic circuits naturally compute a polynomial
- Problem: Different polynomials can give the same function:

$$x^2y + x = xy + x^2 \quad \text{for all } x, y \in \mathbb{F}_2,$$

  but $\mathrm{coeff}_{xy}(x^2y + x) = 0 \neq 1 = \mathrm{coeff}_{xy}(xy + x^2)$.
- The situation is better over infinite fields (for example $\mathbb{C}$):

Lemma

Over an infinite field, two polynomials compute the same function iff they have the same coefficient list.

Proof: Simple induction and polynomial division.

"Algebraic **P** vs **NP**"

The determinant polynomial:

$$\det_m = \sum_{\pi \in \mathfrak{S}_m} \operatorname{sgn}(\pi) \prod_{i=1}^{m} x_{i,\pi(i)}$$

The permanent polynomial:

$$\mathrm{per}_m = \sum_{\pi \in \mathfrak{S}_m} \prod_{i=1}^{m} x_{i,\pi(i)}$$

Assume from now on char $\mathbb{F} \neq 2$, because otherwise $\det_m = \mathrm{per}_m$.

Def.: The **algebraic circuit size** $\mathrm{a}(\mathrm{per}_m)$ is the smallest size of an algebraic circuit computing per_m.

> **Algebraic P vs NP conjecture (VP $\neq$ VNP, Valiant 1979)**
>
> $\mathrm{a}(\mathrm{per}_m)$ is not polynomially bounded.

Determinants instead of circuits

Theorem (Valiant 1979)

Every multivariate polynomial f can be written as the determinant of a matrix whose entries are polynomials of degree ≤ 1.

Example: $$f := y + 2x + xz + 2xy - x^2 z = \det \begin{pmatrix} x & y & 0 \\ -1 & z + y + 2 & x \\ 1 & z & 1 \end{pmatrix}$$

Def.: Required size of the matrix is called the **determinantal complexity** $\mathrm{dc}(f)$.

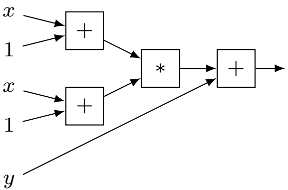
In the example we have $\mathrm{dc}(f) \leq 3$.

Valiant's determinant vs permanent conjecture

$\mathrm{dc}(\mathrm{per}_m)$ is not polynomially bounded.

This is implied by **VP** $\neq$ **VNP**.

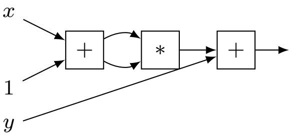
Resources in algebraic computation



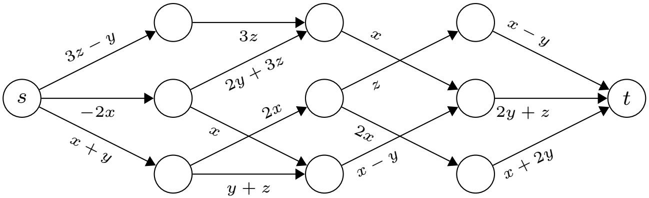
formula size

$$\det\begin{pmatrix} x+1 & y \\ -1 & x+1 \end{pmatrix}$$

determinantal complexity dc



circuit size



- Computes $\sum_{s\text{-}t\text{-path } p} \prod_{\text{edge } e \in p} \text{label}(e)$
- $\mathrm{w}(p)$:= the smallest width of an ABP computing p.

Theorem (Toda 1991)

$\mathrm{dc}(p)$ and $\mathrm{w}(p)$ are polynomially related.

Definition p-family

A **p-family** is a sequence $(f_n)_{n\in\mathbb{N}}$ of polynomials such that:

- The number of variables is polynomially bounded
- The degree is polynomially bounded

- **VF** := { p-family whose formula size is polynomially bounded }
- **VBP** := { p-family whose dc (or w) is polynomially bounded }
- **VP** := { p-family whose circuit size is polynomially bounded }

VF $\subseteq$ **VBP** $\subseteq$ **VP**

$(f_n) \in C$ is complete for C if $\forall (g_m) \in C$ there exists a polynomially bounded s and linear polynomials ℓ_i such that

$$\forall m: \quad g_m = f_{s(m)}(\ell_1, \ell_2, \ldots).$$

For example, $(\det_n)$ is **VBP**-complete.
Example: $(x_1 x_2 \cdots x_n) \in$ **VBP**, because $\det(\text{diag}(x_1, x_2, \ldots, x_n)) = x_1 x_2 \cdots x_n$.

$$\mathrm{IMM}_r^{(d)} := \begin{pmatrix} x_{1,1,1} & x_{1,2,1} & \cdots & x_{1,r,1} \end{pmatrix} \begin{pmatrix} x_{1,1,2} & \cdots & x_{1,r,2} \\ \vdots & \ddots & \vdots \\ x_{r,1,2} & \cdots & x_{r,r,2} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1,d-1} & \cdots & x_{1,r,d-1} \\ \vdots & \ddots & \vdots \\ x_{r,1,d-1} & \cdots & x_{r,r,d-1} \end{pmatrix} \begin{pmatrix} x_{1,1,d} \\ \vdots \\ x_{1,r,d} \end{pmatrix}$$

- $\mathrm{IMM}_3^{(n)}$ is **VF**-complete [Ben-Or, Cleve 1988].
- $\mathrm{IMM}_n^{(n)}$ is **VBP**-complete.
- There is no equally nice **VP**-complete p-family known.

Definition **VNP**

A p-family (f_n) is in **VNP** if there exists a p-family $(g_n) \in$ **VP** and polynomially bounded functions r, s, t such that

$$\forall n: \ f_n = \sum_{b \in \{0,1\}^{r(n)}} g_{t(n)}(x_1, \ldots, x_{s(n)}, b_1, \ldots, b_{r(n)})$$

For example,

$$\mathrm{per}_n = \sum_{b \in \{0,1\}^{n^2}} C(b) \prod_{1 \le i,j \le n} (b_{i,j}(x_{i,j} - 1) + 1),$$

where C is the arithmetization of a Boolean circuit checking if b is a permutation matrix.

- One can also take $(g_n) \in$ **VF** and it gives the same class: **VNP** $=$ **VNF**.
- One can also take (g_n) to be just a polynomially long product of linear polynomials (Bringmann-I-Zuiddam 2018)

Valiant 1979:

- The permanent p-family (per_n) is **VNP**-complete.

Efficient computation:

- $\mathbf{VF} \subseteq \mathbf{VBP} \subseteq \mathbf{VP}$
- "$\mathbf{VBP}$ = linear algebra" (determinant, iterated matrix multiplication)

Efficiently definable ("explicit polynomials"):

- **VNP**
- "$\mathbf{VNP}$ = combinatorics/counting" (Cycle covers, permanent)

Valiant's conjectures

$\mathbf{VF} \neq \mathbf{VNP}$

$\mathbf{VBP} \neq \mathbf{VNP}$, determinant vs permanent, linear algebra vs counting

$\mathbf{VP} \neq \mathbf{VNP}$

Valiant's conjectures

$\mathbf{VF} \neq \mathbf{VNP}$

$\mathbf{VBP} \neq \mathbf{VNP}$, determinant vs permanent, linear algebra vs counting

$\mathbf{VP} \neq \mathbf{VNP}$

These algebraic conjectures are "easier" than the Boolean ones:

$$\mathbf{PH} \neq \Sigma_2 \overset{\text{Karp-Lipton 1982}}{\Longrightarrow} \mathbf{NP} \not\subseteq \mathbf{P/poly} \overset{\text{Bürgisser 2000}}{\Longrightarrow} \mathbf{VP} \neq \mathbf{VNP} \implies \mathbf{VBP} \neq \mathbf{VNP} \implies \mathbf{VF} \neq \mathbf{VNP}$$

Bürgisser's result works

- over finite fields, and
- over $\mathbb{C}$ (if the generalized Riemann hypothesis is true).

Agenda

We work with homogenized algebraic branching programs!
In fact, inhomogeneous set-ups can lead to weird behavior in the representation theory (Landsberg-Kadish 2012, I-Panova 2015, Bürgisser-I-Panova 2015)

For a fixed degree d and number of variables n and a complexity bound r, study the set

$$X_r := \{f \in \mathbb{C}[x_1, \dots, x_n]_d \mid \mathrm{w}(f) \leq r\}.$$

$$X_1 \subseteq X_2 \subseteq X_3 \subseteq \cdots \subseteq X_{\max} = \mathbb{C}[x_1, \dots, x_n]_d$$

$$\mathrm{IMM}_r^{(d)} := \begin{pmatrix} x_{1,1,1} & x_{1,2,1} & \cdots & x_{1,r,1} \end{pmatrix} \begin{pmatrix} x_{1,1,2} & \cdots & x_{1,r,2} \\ \vdots & \ddots & \vdots \\ x_{r,1,2} & \cdots & x_{r,r,2} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1,d-1} & \cdots & x_{1,r,d-1} \\ \vdots & \ddots & \vdots \\ x_{r,1,d-1} & \cdots & x_{r,r,d-1} \end{pmatrix} \begin{pmatrix} x_{1,1,d} \\ \vdots \\ x_{1,r,d} \end{pmatrix}$$

VBP completeness with homogenization gives:

- If $f \in X_r$, then $f(A\vec{x}) \in X_r$ for any linear map A.
- Every $f \in X_r$ can be obtained via a linear map A as $f = \mathrm{IMM}_r^{(d)}(A\vec{x})$

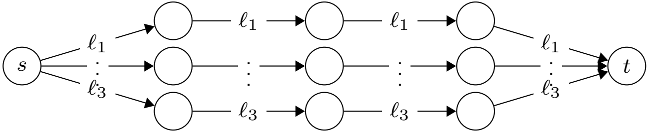
For example, if $f = x_1^3 + x_1x_2x_3$, and $A = \left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{smallmatrix}\right)$, then $f(A\vec{x}) = f(x_2, x_1, x_1 + x_2) = x_2^3 + x_1^2x_2 + x_1x_2^2$.

$$X_r := \{f \in \mathbb{C}[x_1, \dots, x_n]_d \mid \mathrm{w}(f) \le r\}.$$

- Goal: find more useful mathematical structure on X_r.
- X_r is closed under base changes: Changing input variables to linear combination comes at no extra cost.
- But X_r is lacking a crucial property: It is not topologically closed.

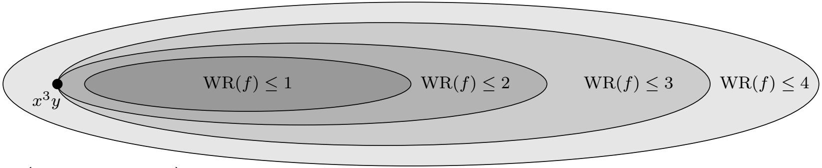
Example from now on via Waring rank WR instead of w.

For a homogeneous degree d polynomial f the **Waring rank** $\mathrm{WR}(f)$ is defined as the smallest r such that there exist homogeneous linear ℓ_i such that $f = \sum_{i=1}^{r} (\ell_i)^d$.



Not entirely obvious at first: $\mathrm{WR}(f)$ is always finite.
For example:
$12x^3y = (x+y)^4 + i^3(x+iy)^4 + i^2(x+i^2y)^4 + i(x+i^3y)^4$, hence $\mathrm{WR}(x^3y) \le 4$. In fact, $\mathrm{WR}(x^{d-1}y) = d$.

$12x^3y = (x+y)^4 + i^3(x+iy)^4 + i^2(x+i^2y)^4 + i(x+i^3y)^4.$



$\frac{1}{\varepsilon}\Big((x+\varepsilon y)^4 - x^4\Big) = 4x^3y + \varepsilon(6x^2y^2 + 4\varepsilon xy^3 + \varepsilon^2 y^4) \xrightarrow{\varepsilon\to 0} 4x^3y$

The **border Waring rank** $\underline{\mathrm{WR}}(f)$ is defined as the smallest r such that f can be approximated arbitrarily closely by polynomials of Waring rank $\leq r$. For example, $\underline{\mathrm{WR}}(x^3y) \leq 2$.

For a space of polynomials $V = \mathbb{C}[x_1,\ldots,x_n]_d$, the elements of $\mathbb{C}[V]$ are called **meta-polynomials**.

Example: For $ax^2 + bxy + cy^2$, the discriminant $b^2 - 4ac$ is a meta-polynomial.

Theorem (algebraic geometry)

The set $\overline{W_r} = \{f \mid \underline{\mathrm{WR}}(f) \leq k\}$ is an algebraic variety, i.e., there exist finitely many meta-polynomials $\Delta_1,\ldots,\Delta_N$ with

$$f \in \overline{W_r} \quad \Leftrightarrow \quad \Delta_1(f) = \Delta_2(f) = \cdots = \Delta_N(f) = 0$$

We conclude that we know how complexity lower bounds must look like:

Theorem: If $f \notin \overline{W_r}$, then there exists a meta-polynomial Δ with • $\Delta(\overline{W_r}) = \{0\}$ and • $\Delta(f) \neq 0$

$\mathbb{A} := \mathbb{C}[x,y]_2 = \langle x^2, xy, y^2 \rangle$.

Every element in $\mathbb{A}$ can be represented as $ax^2 + bxy + cy^2$.

- $X := \{f \in \mathbb{A} \mid \exists \alpha, \beta \in \mathbb{C} : f = (\alpha x + \beta y)^2\}$ set of Waring rank 1 polynomials
- $f \in X$ iff $\Delta(f) = b^2 - 4ac = 0$.
- To prove $\mathrm{WR}(f) \geq 2$ we compute $\Delta(f) \neq 0$

Topological closures of algebraic complexity classes

$\mathrm{WR} \quad \longrightarrow \quad \underline{\mathrm{WR}}$

Analogously, we can allow such approximations

- for formulas,
- for algebraic branching programs,
- for circuits,
- or for hypercube summations of circuits.

The corresponding complexity classes are

- $\mathbf{VF} \subseteq \overline{\mathbf{VF}}$,
- $\mathbf{VBP} \subseteq \overline{\mathbf{VBP}}$,
- $\mathbf{VP} \subseteq \overline{\mathbf{VP}}$,
- $\mathbf{VNP} \subseteq \overline{\mathbf{VNP}}$.

Wide open question: Is $\overline{\mathbf{VF}} \subseteq \mathbf{VNP}$?

Let $X_r := \{f \in \mathbb{C}[x_1, \dots, x_n]_d \mid \mathrm{w}(f) \le r\}$.

- To find meta-polynomials Δ for $\overline{X_r}$, instead of studying $\overline{X_r}$ directly, one can study its **coordinate ring**, i.e., ring of polynomial funtions restricted to $\overline{X_r}$.
- Representation theory helps to study the coordinate ring using the weighs of the GL_n: a finer variant of a degree for meta-polynomials
- Connections to invariant theory and algebraic combinatorics, established by Mulmuley and Sohoni 2001, 2008.

Conclusion

- We do not know how powerful approximations in algebraic complexity theory are, but
- if we allow approximations, then all complexity lower bounds come from meta-polynomials, and this opens a wide array of tools from algebraic geometry and representation theory.

Thank you very much for your attention!