Lecture 2: Algebraic Circuits & Algebraic Complexity

Rafael Oliveira

University of Waterloo
Cheriton School of Computer Science

rafael.oliveira.teaching@gmail.com

January 13, 2021

Overview

- Algebraic Complexity Classes
- Structural Results on Algebraic Circuits
- Conclusion
- Acknowledgements

Complexity Measures in Algebraic Circuits

- *circuit size:* number of edges in the circuit, denoted by $\mathcal{S}(\Phi)$
- *cost of ring elements:* in classical algebraic complexity, there is unit cost for the use of any base ring element
- Sometimes we will add bit complexity of base ring elements

$P(x,y) = x^2 + 2xy - 3y + 1$

sparse reprentation

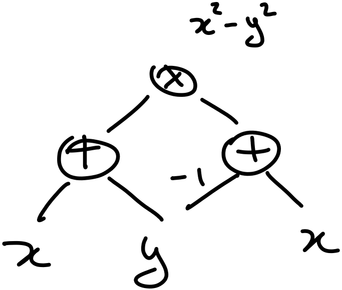
$(1, x^2)$, $(2, xy)$, $(-3, y)$, $(1, 1)$

dense representation

$[2, (1, x^2), (2, xy), (0, y^2), (0, x), (-3, y), (1, 1)]$

Complexity Measures in Algebraic Circuits

- *circuit size:* number of edges in the circuit, denoted by $\mathcal{S}(\Phi)$
- *cost of ring elements:* in classical algebraic complexity, there is unit cost for the use of any base ring element
- Sometimes we will add bit complexity of base ring elements
- *circuit depth:* length of longest direct path from an input to an output



depth-2

depth = parallel complexity

Complexity Measures in Algebraic Circuits

- *circuit size:* number of edges in the circuit, denoted by $\mathcal{S}(\Phi)$
- *cost of ring elements:* in classical algebraic complexity, there is unit cost for the use of any base ring element
- Sometimes we will add bit complexity of base ring elements
- *circuit depth:* length of longest direct path from an input to an output
- *constant depth circuits:* for circuits of constant depth, we don't place restriction on the fan-in of an edge.

Convention: whenever we don't specify depth, the fanin of each gate is ≤ 2

circuits of constant depth may have arbitrary fanin.

Examples - Constant Depth Circuits

$\Sigma\Pi$ - ckts $\longleftrightarrow$ sparse polynomials

$\prod_{i=1}^{n}(x_i+1)$ $\quad 2^n$ size $\Sigma\Pi$ - ckts

$\Sigma\Pi\Sigma$ -ckts $\longleftrightarrow$ $\sum_{i=1}^{s}\prod_{j=1}^{d} \ell_{ij}(x_1,\dots,x_n)$

~~$\Pi\Sigma\Pi$ prod. of sparse poly~~

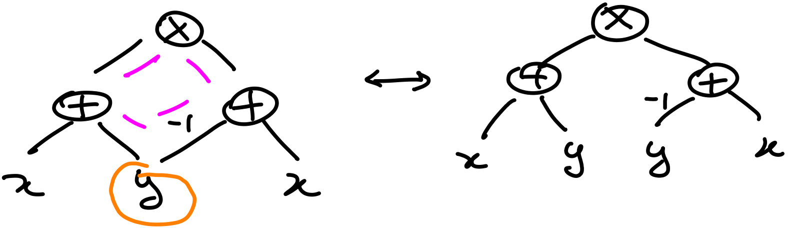
$\prod_{i=1}^{n}(x_i+1)$ $\quad O(n)$ size $\Sigma\Pi\Sigma$ - ckt

$\Sigma\Pi\underbrace{\Sigma\Pi}$ - ckts $\longleftrightarrow$ $\sum_{i=1}^{s}\prod_{j=1}^{d} \overbrace{P_{ij}}^{\Sigma\Pi}$

P_{ij} sparse polynomials

Algebraic Formulas

- when the computation graph is a tree (i.e., we don't reuse computation) we get an algebraic formula



Result: P has ckt of size s, then

P has a formula of size $s^{\log d \log n} = s^{\log^2 n}$

poly-size ckt $\subset$ quasi-poly formula

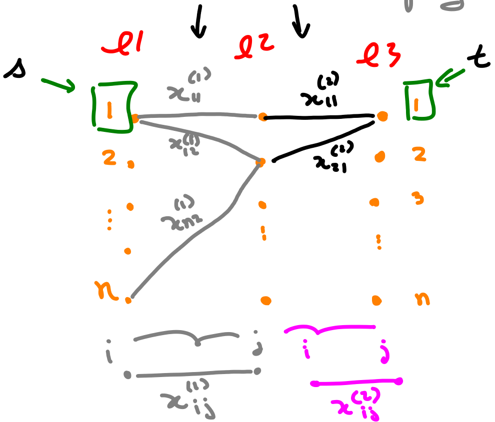
Algebraic Branching Programs

- polynomials which are projections of the *Iterated Matrix Multiplication* (IMM) polynomial

$$X_i = \left(x^{(i)}_{jk}\right)^{n}_{j,k=1}$$

$$\underbrace{\mathrm{tr}[X_1 X_2 \cdots X_d]}_{\text{poly in } n^2 \cdot d \text{ variables}}$$

d degree

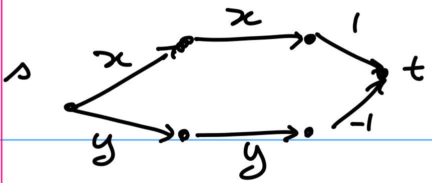


$$P_{1,1} = \sum_k x^{(1)}_{1k} \cdot x^{(2)}_{k1}$$

$$[X_1 X_2]_{11} = P_{11}$$

$$P = X_1 X_2$$

$$P_{ij}$$



expressed as $t(X_1 X_2 X_3)$

$$\sum_{s-t \text{ path}} \prod (\text{vars in path})$$

$$x \cdot x \cdot 1 + y \cdot y \cdot (-1) = x^2 - y^2$$

Algebraic branching program

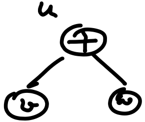
Homogeneous Components

Theorem ([Strassen 1973])

If a polynomial $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ can be computed by a circuit Φ of size $\mathcal{S}(\Phi)$, then the homogeneous components $H_0[p], H_1[p], \ldots, H_r[p]$ can be computed by a homogeneous circuit of size $O(r^2 \cdot \mathcal{S}(\Phi))$.

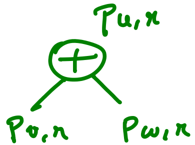
Proof: induction over depth.

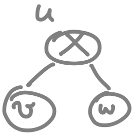
Base: input gates homogeneous.



$p_u = p_v + p_w$

$H_r[p_u] = H_r[p_v] + H_r[p_w]$

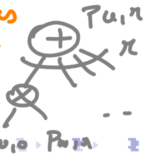




$p_u = p_v \cdot p_w$

$$H_r[p_u] = \sum_{d=0}^{r} \underbrace{H_d[p_v]}_{d} \cdot \underbrace{H_{r-d}[p_w]}_{r-d}$$

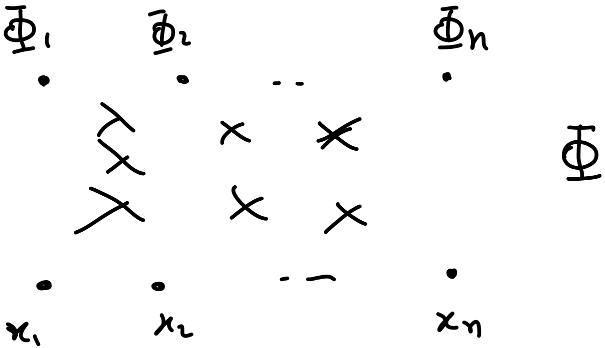
$O(\log r)$ r gates ←



Universal Circuits

Definition

A circuit Φ is called (n, d, s)-*universal*, if the following holds:
If $f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)$ are homogeneous polynomials of degree d which can be simultaneously computed by a circuit of size s, then there is a circuit $\boxed{\Psi}$ computing $f_1, \ldots, f_n$ with same computation graph as $\boxed{\Phi}$.



Universal Circuits

Definition

A circuit Φ is called (n, d, s)-*universal*, if the following holds:
If $f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)$ are homogeneous polynomials of degree d which can be simultaneously computed by a circuit of size s, then there is a circuit Ψ computing $f_1, \ldots, f_n$ with same computation graph as Φ.

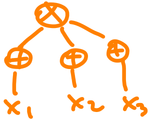
- Φ is (n, d, s)-universal if any circuit Ψ of size $\leq s$ computing homogeneous polynomials of degree d are a projection of Φ

Universal Circuits

Definition

A circuit Φ is called (n, d, s)-*universal*, if the following holds:
If $f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)$ are homogeneous polynomials of degree d which can be simultaneously computed by a circuit of size s, then there is a circuit Ψ computing $f_1, \ldots, f_n$ with same computation graph as Φ.

- Φ is (n, d, s)-universal if any circuit Ψ of size $\leq s$ computing homogeneous polynomials of degree d are a projection of Φ
- Normal-homogeneous form:
  - all input gates are labelled by a variable
  - all edges leaving input gates are connected to sum gates
  - all output gates are sum gates
  - alternating sum-product layers
  - fanin of each *product* gate is exactly 2
  - out-degree of each *addition* gate is ≤ 1



formula like for addition

$\Sigma \, \Pi \, \Sigma \, \Pi \, \Sigma \, \Pi \, \Sigma$

Universal Circuits

Theorem ([Raz 2008])

For any integers $s \geq n$ and d, we can construct in time $\mathrm{poly}(s, d)$ a circuit Φ in normal-homogeneous form with at most $O(s^4 d)$ nodes that is (n, d, s)-universal.

$s \cdot d^4$

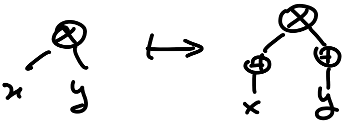
- For every circuit Ψ, there is a circuit χ in normal homogeneous form computing all polynomials that Ψ computes

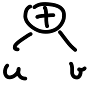
$O(s \cdot d^2)$

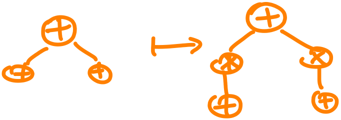
s

$\Psi \longmapsto \underbrace{H_d[\Psi]}_{\text{homogeneous}}$

P $\qquad$ P





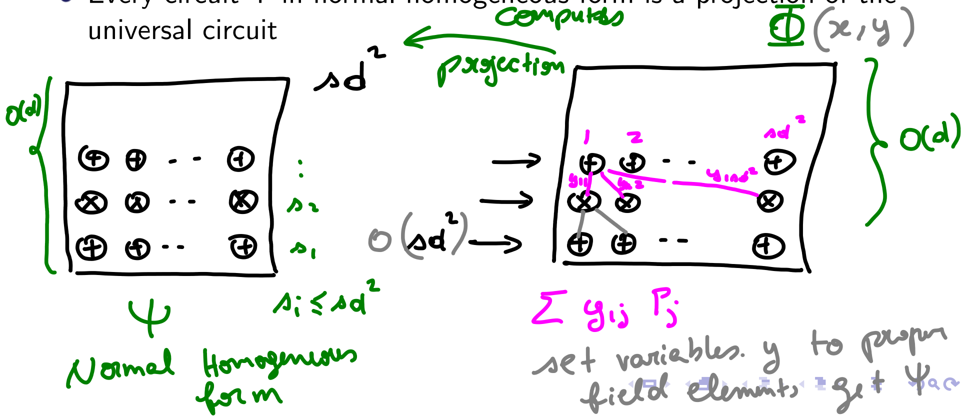


Universal Circuits

Theorem ([Raz 2008])

For any integers $s \geq n$ and d, we can construct in time $\mathrm{poly}(s, d)$ a circuit Φ in normal-homogeneous form with at most $O(s^4 d)$ nodes that is (n, d, s)-universal.

- Every circuit Ψ in normal homogeneous form is a projection of the universal circuit



Almost all polynomials are hard to compute

Corollary: the set of polynomials which can be computed by small circuits has "measure zero" over $\mathbb{F}[x_1, .., x_n]$.

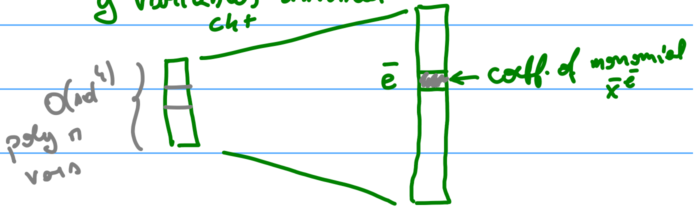
Proof: look $\mathbb{F}[x_1, .., x_n]_d$

dim $\binom{n+d-1}{d}$ d^n or n^d

if polynomial $f \in \mathbb{F}[x_1, .., x_n]_d$ computed by ckt of size s then f can be computed by projection of universal ckt $O(sd^4)$ $(\mathrm{poly}(n))$

$$\Phi : \underbrace{\mathbb{F}^{O(sd^4)}}_{y \text{ variables universal ckt}} \longrightarrow \mathbb{F}[x_1,\dots,x_n]_d$$

$g_{\bar{e}}(y_1,\dots,y_{O(sd^4)})$



universal ckt

$$\mathrm{Im}(\Phi) \supset \left\{ f \;\middle|\; \begin{array}{c} f \text{ computed} \\ \text{by small} \\ \text{ckts} \end{array} \right\}$$

$$\dim\left(\overline{\mathrm{Im}(\Phi)}\right) \le O(sd^4)$$

$$\ll \binom{n+d-1}{d}$$

$\Rightarrow$ most polys hard.

Computing Partial Derivatives

$\Phi^{w} \to \Phi^{w}_{y}$

Theorem ([Baur, Strassen 1983])

If a polynomial $p(x_1, \dots, x_n) \in \mathbb{F}[x_1, \dots, x_n]$ can be computed by a circuit Φ of size s and depth d, then there is a circuit Ψ of size $O(s)$ and depth $O(d)$ computing (simultaneously) the polynomials $\partial_1 p, \partial_2 p, \dots, \partial_n p$.

Notation: v, w two gates in Φ

computing $f_v, f_w \in \mathbb{F}[x_1, \dots, x_n]$

if delete wires into w (make it "input") and label it by new variable y, then

- $\Phi_{w=y}$ is our new ckt
- $f_{v,w}(x_1, \dots, x_n, y)$ polynomial computed in gate v in $\Phi_{w=y}$.

Computing Partial Derivatives

Theorem ([Baur, Strassen 1983])

If a polynomial $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ can be computed by a circuit Φ of size s and depth d, then there is a circuit Ψ of size $O(s)$ and depth $O(d)$ computing (simultaneously) the polynomials $\partial_1 p, \partial_2 p, \ldots, \partial_n p$.

- Taking partial derivative with respect to a gate

$$\underbrace{\partial_w f_v}_{\in \mathbb{F}[x_1, \ldots, x_n]} := \left(\partial_y \underbrace{f_{v,w}}_{} \right)\Big|_{\underline{y = f_w}}$$

partial derivative w.r.t. w of gate v

w input gate with variable y

substitute y by f_w

Computing Partial Derivatives

Φ original

Ψ partial derivatives

- Induction on circuit size

Base case: input gates ✓

Induction ckt size: Φ let w be the lowest gate that we have not computed $\Psi_{\partial_1 f_w} \dots \Psi_{\partial_n f_w}$

$S(\Phi_{w=y}) \leq S(\Phi) - 1$ (removed two edges)

By induction can compute $\Psi_{w=y}$

Computing Partial Derivatives

Chain rule:

$$\partial_i f_v = \underbrace{\partial_i f_{v,w} \Big|_{y=f_w}}_{\text{in } \Psi_{w=y}} + \underbrace{\partial_y f_{v,w} \Big|_{y=f_w}}_{\text{in } \Psi_{w=y}} \cdot \underbrace{\partial_i f_w}_{\text{in } \Psi_{w=y}}$$

(because Φ be subckt over Ψ $S(\Psi) + S(\Phi)$)

compute $\partial_i f$ by adding constant number of edges.

Computing Partial Derivatives

Open problem: can we also compute second order partial derivatives with only a constant blow-up in ckt size?

non-trivial

Consequence: if yes, we would get that matrix multiplication can be computed in $O(n^2)$ time.

nontrivial

Practice problem: prove this consequence.

Depth Reduction

Theorem ([Valiant, Skyum, Berkowitz, Rackoff 1983])

If a homogeneous polynomial $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ of degree d can be computed by a circuit of size s, there is a homogeneous circuit Φ of size $\text{poly}(d, s)$ computing p such that:

1. *Φ has alternating levels of sum and product gates*
2. *each product gate $v \in \Phi$ computes the product of five polynomials, each of degree $\leq 2 \cdot \deg(v)/3$*
3. *Sum gates have arbitrary fanin* $\}$ $O(\log s)$ depth fanin 2

In particular, the number of levels in Φ is $O(\log d)$.

The above yields circuit (with fanin 2 for every gate) of depth

$$O(\underbrace{\log d}_{\#\text{levels}}\underbrace{(\log d + \log s)}_{\hookrightarrow \text{ sum gates fanin 2}})$$

Depth Reduction

$NC^1 \subset NC^2 \subset \cdots$

Theorem ([Valiant, Skyum, Berkowitz, Rackoff 1983])

If a homogeneous polynomial $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ *of degree* d *can be computed by a circuit of size* s*, there is a homogeneous circuit* Φ *of size* $\mathrm{poly}(d, s)$ *computing* p *such that:*

1. Φ *has alternating levels of sum and product gates*
2. *each product gate* $v \in \Phi$ *computes the product of five polynomials, each of degree* $\leq 2 \cdot \deg(v)/3$
3. *Sum gates have arbitrary fanin*

In particular, the number of levels in Φ *is* $O(\log d)$*.*

The above yields circuit (with fanin 2 for every gate) of depth

$$O(\log d(\log d + \log s))$$

Corollary

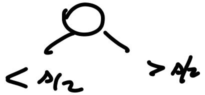
$$VP = VNC^2$$

Depth Reduction

Theorem (Depth Reduction for Formulas)

If $p(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ is computed by a formula of size s, then $p(x_1, \ldots, x_n)$ can also be computed by a formula of depth $O(\log s)$.

Depth Reduction



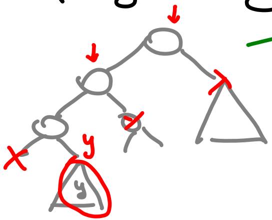
Φ size s

find gate v s.t. $S(\Phi_v)$ satisfies

$$\frac{s}{3} \le S(\Phi_v) \le \frac{2s}{3}$$

$\Phi_{v=y}$

$$f_{v=y} = \underbrace{y \cdot \partial_v f}_{\text{formula}} + f_{v=y}\Big|_{y=0}$$



$$f = f_v \cdot \partial_v f + \overbrace{f_{v=y}\big|_{y=0}}^{s - S(\Phi_v)}$$

Depth Reduction

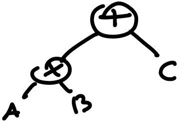
$$\frac{s}{3} \le S(\Phi_v) \le \frac{2s}{3}$$

$$\left(\frac{2}{3}\right)^i s \to O(i)$$

$$i = O(\log s)$$

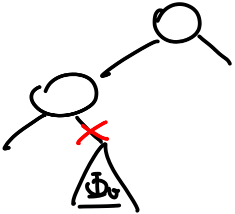
$$\underbrace{s - S(\Phi_v)}_{f_{v=y}|_{y=0}} \le \frac{2s}{3}$$

$$f = \overbrace{f_v}^{\le \frac{2s}{3}} \cdot \underbrace{\partial_v f}_{\le \frac{2s}{3}} + f_{v=y}\big|_{y=0}$$



$$\Phi = A \cdot B + C$$

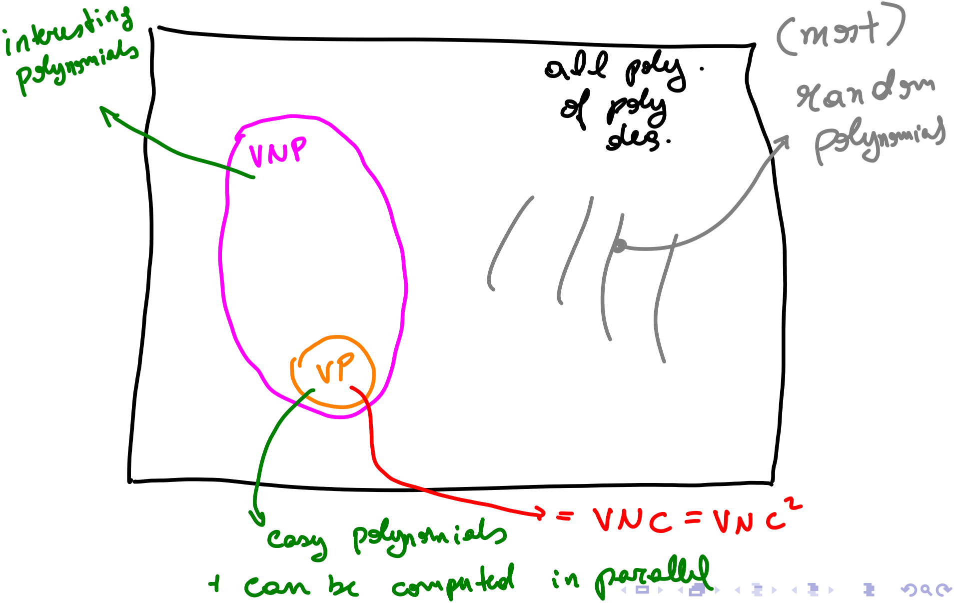
$$s(A),\ s(B),\ s(C) \le \frac{2s}{3}$$



Depth Reduction

$$VP_{formulas} \underset{\supset}{\subset} VNC^1$$

Our world map so far



Conclusion

- Today we learned some additional algebraic complexity classes
  - constant depth circuits (formulas)
  - algebraic branching programs
  - algebraic formulas
- Construction of universal circuit
- Efficient computation of partial derivatives using algebraic circuits
- Depth reduction
- Consequences to algebraic complexity

Acknowledgement

- Lecture based largely on:
  - Excellent survey [Shpilka & Yehudayoff 2010, Chapter 2]

    https://www.nowpublishers.com/article/Details/TCS-039

References I

Valiant, L. and Skyum, S. and Berkowitz, S. and Rackoff, C. 1983.
Fast parallel computation of polynomials using few processors
SIAM Journal on Computing

Baur, W. and Strassen, V. 1983.
The complexity of partial derivatives
Theoretical Computer Science

Shpilka, Amir and Yehudayoff, Amir 1982.
Arithmetic circuits: a survey of recent results and open questions
Foundations and Trends in Theoretical Computer Science

Raz, Ran 2008.
Elusive functions and lower bounds for arithmetic circuits
STOC

Strassen, Volker 1973.
Vermeidung von Divisionen
The Journal fur die Reine und Angewandte Mathematik