Upper bounds on real roots and lower bounds for the permanent

Pascal Koiran
LIP, Ecole Normale Supérieure de Lyon

ISSAC 2012 Tutorial
Grenoble, July 22, 2012

The material:

- Upper bounds on number of real roots for certain sparse polynomial systems.
- Depth reduction for arithmetic circuits.

The motivating problem:
What is the arithmetic complexity of the permanent polynomial?
This is:

- An arithmetic version of P=NP (Valiant'79).
- Roughly equivalent to determinant versus permanent.

Reminder: $\mathrm{per}(X) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_{i\sigma(i)}$.

Determinant versus permanent (1/2)

Representing a permanent by a determinant:

$$\text{per}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det\begin{bmatrix} a & -b \\ c & d \end{bmatrix}$$

$$\text{per}\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \det\begin{bmatrix} 0 & a & d & g & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & i & f & 0 \\ 0 & 0 & 1 & 0 & 0 & c & i \\ 0 & 0 & 0 & 1 & c & 0 & f \\ e & 0 & 0 & 0 & 1 & 0 & 0 \\ h & 0 & 0 & 0 & 0 & 1 & 0 \\ b & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The general case: A permanent of size n can be represented by a determinant of size $2^n - 1$ (B. Grenet).

Determinant versus permanent (2/2)

Conjecture:
If $\text{per}(A) = \det(B)$ then $\mathrm{size}(B)$ cannot be polynomial in $\mathrm{size}(A)$.
The entries of B can be either:

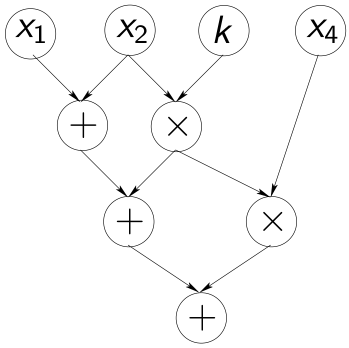
- Entries of A, or constants.
- Affine functions of the entries of A.

Remark: These 2 versions of the conjecture are equivalent:
det(affine functions) $\rightarrow$ det(variables or constants).

Some work toward the conjecture:

- $\mathrm{size}(B) \geq \mathrm{size}(A)^2/2$ (Mignon and Ressayre, 2004).
- Geometric Complexity Theory:
  an approach based on representation theory
  (Ketan Mulmuley / Milind Sohoni + Bürgisser, Kumar, Landsberg, Manivel, Ressayre, Weyman...).
- Today's approach is based on sparse polynomials,
  and uses the completeness of the permanent.

Arithmetic circuits: Toward an arithmetic version of P versus NP



Circuit
Size: 9
Depth: 3

Valiant's model: $VP_K = VNP_K$?

- Complexity of a polynomial f measured by number $L(f)$ of arithmetic operations (+,-,$\times$) needed to evaluate f:

  L(f) = size of smallest arithmetic circuit computing f.
- $(f_n) \in VP$ if number of variables, $\deg(f_n)$ and $L(f_n)$ are polynomially bounded.
  Two examples: the determinant family $(\det_n)$ is in VP, but $(X^{2^n}) \notin VP$.
- $(f_n) \in VNP$ if $f_n(\overline{x}) = \sum_{\overline{y}} g_n(\overline{x}, \overline{y})$

  for some $(g_n) \in VP$
  (sum ranges over all boolean values of $\overline{y}$).
  Example:
  If $\mathrm{char}(K) \neq 2$ the permanent is a VNP-complete family.

Overview of the tutorial

1. Depth reduction for arithmetic circuits:
   - Reduction to depth $O(\log n)$ for arithmetic formulas (Muller-Preparata'76).
   - Reduction to depth $O(\log^2 n)$ for low-degree circuits (Valiant-Skyum-Berkowitz-Rackoff'83).
   - **Reduction to depth 4 for low-degree circuits** (Agrawal-Vinay, 2008).
2. The real τ-conjecture:
   a connection between sparse polynomials
   and lower bounds for the permanent.
3. Upper bound on the number of real roots.

Sparse polynomials: a glimpse of part 3

- Descartes' rule without signs:
  If f has t monomials then f at most $t-1$ positive real roots.
- Khovanskii's theory of fewnomials: a system

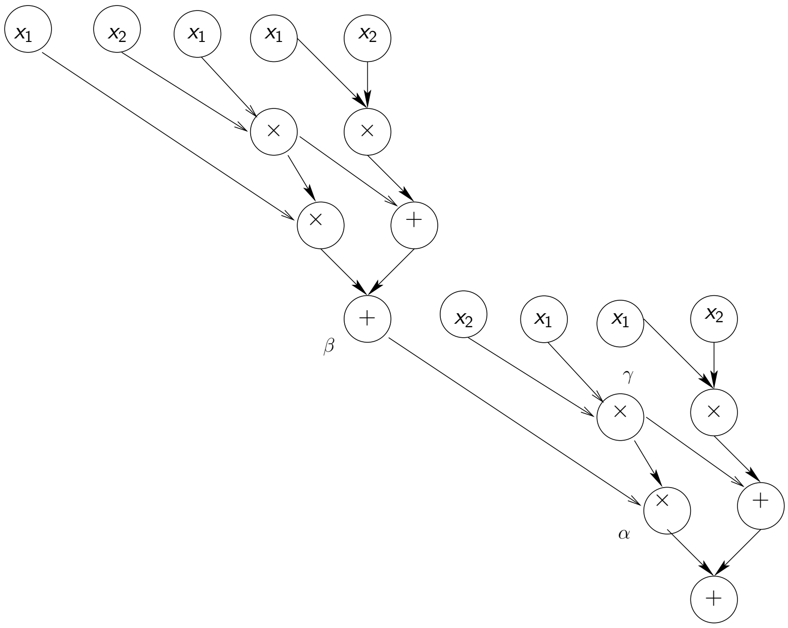
  $$f_1(x_1,\ldots,x_n) = f_2(x_1,\ldots,x_n) = \cdots = f_n(x_1,\ldots,x_n) = 0$$

  with t distinct exponent vectors has at most $(n+1)^t 2^{t(t-1)/2}$ non-degenerate roots in the positive orthant.
- For certain sparse systems,
  the *Wronskian determinant* leads to better bounds.

A take-home problem:
How many real solutions to the univariate equation $fg = 1$?
Descartes' bound is $O(t^2)$ but true bound could be $O(t)$.
Remark: $fg = 1$ can be re-written as $[y = f(x), y.g(x) = 0]$.

Weakly Skew Circuits

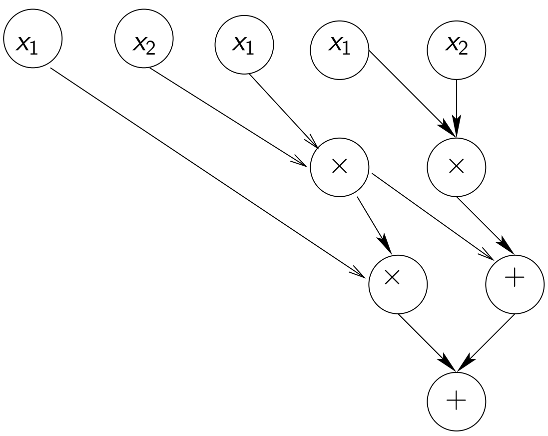
For each multiplication gate $\alpha := \beta \times \gamma$:
C_β or C_γ is independent from the remainder of the circuit.



If a gate is not in an independent subcircuit it is *reusable*.

Skew Circuits

For each multiplication gate $\alpha := \beta \times \gamma$:
β or γ is an input.



Skew Circuits $\subseteq$ Weakly Skew Circuits,
and Arithmetic Formulas (Trees) $\subseteq$ Weakly Skew Circuits.

(Weakly) Skew Circuits and the Determinant

Weakly skew circuits capture the complexity of the determinant.

Theorem (Toda92)

The determinant can be computed by:

- *Weakly skew circuits of size $O(n^7)$.*
- *Skew circuits of size $O(n^{20})$.*

Proof based on Berkowitz's algorithm.

Theorem (Toda92,Malod03)

A weakly skew circuit of size t has an equivalent determinant (and permanent) of size $t+1$.

Applications

- Closure properties of the determinant:
  1. Stability under polynomial size summation [Malod - Portier'06-08]
  2. Stability under exact quotient [Kaltofen - Koiran'08]
  3. det(affine functions) $\rightarrow$ det(variables or constants).

  Proof: convert determinants into weakly skew circuits, convert back final result into determinant form.
- Expressive power of determinants of symmetric matrices [Grenet-Kaltofen-Koiran-Portier'11]

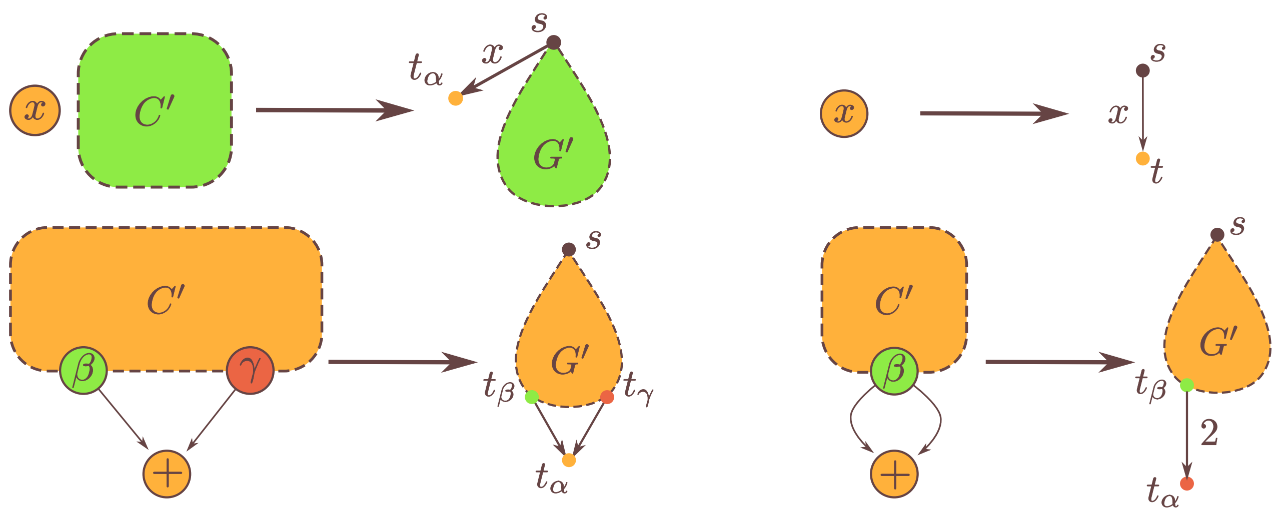
From Weakly Skew Circuit to Determinants (1/4)

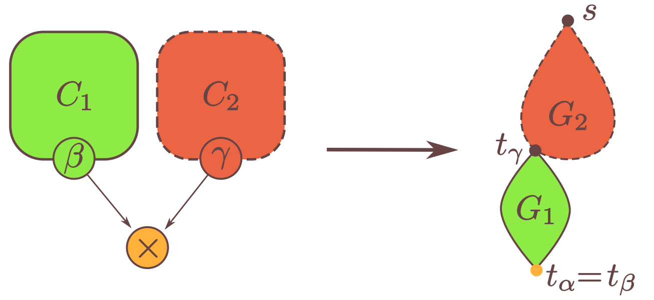
An arithmetic branching programs is a dag with two distinguished vertices s, t.

- edges labeled by variables or constants.
- weight of path = product of edge weights.
- output = $w(s \to t)$ = sum of the weights of all st-paths.

(Valiant'79, universality of per/det for arithmetic formulas.)

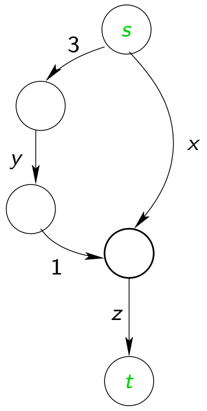
From Weakly Skew Circuit to Determinants (2/4)

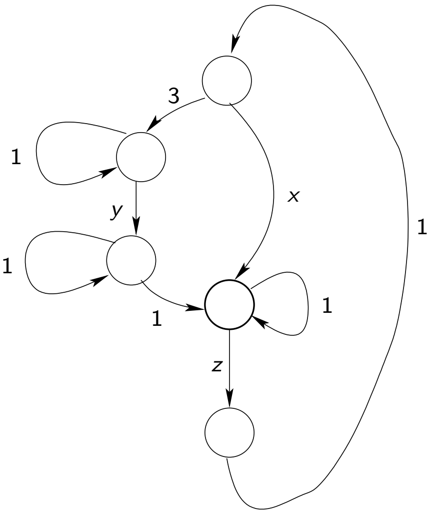




Invariant:
For each *reusable* gate α, there exists t_α s.t. $w(s \to t_\alpha) = \phi_\alpha$.

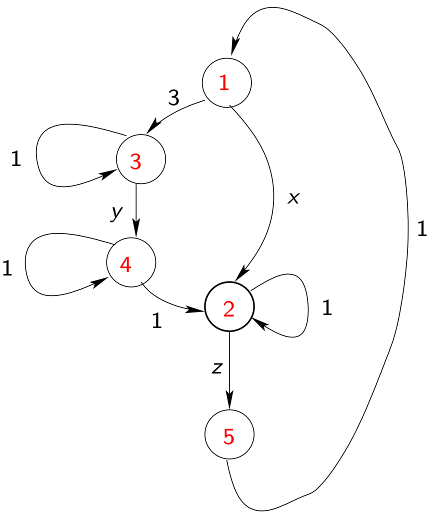
From Weakly Skew Circuit to Determinants (3/4)





From Weakly Skew Circuit to Determinants (4/4)

$$\det \begin{pmatrix} 0 & x & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 & z \\ 0 & 0 & 1 & y & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$



$$\mathrm{per} A = \sum_{\sigma} \prod_{i=1}^{n} A_{i,\sigma(i)}; \quad \det A = \sum_{\sigma} (-1)^{\mathrm{sgn}(\sigma)} \prod_{i=1}^{n} A_{i,\sigma(i)}$$

Permutation in A = cycle cover in G.

Up to signs, $\det A$ = sum of weights of cycle covers in G.

More on Skew versus Weakly Skew

Theorem (Kaltofen-Koiran'08, Jansen'08)

A weakly skew circuit of size m has an equivalent skew circuit of size 2m.

1. Construct equivalent arithmetic branching program G of size $m+1$.
2. Compute inductively $w(s \to v)$ for each node $v \in G$.
   - Two predecessors v_1, v_2 with unit edge weights: $w(s \to v) = w(s \to v_1) + w(s \to v_2)$.
   - One predecessor v_1 with edge weight x: $w(s \to v) = x \times w(s \to v_1)$.

Parallelization of Weakly Skew Circuits

Theorem: Let G be an branching program of size m and depth δ. There is an equivalent circuit of depth $2\log\delta$, with $m^3 \log\delta$ binary multiplication gates and $m^2 \log\delta$ addition gates of unbounded fan-in.

Consequence: polynomial size weakly skew circuits
$\Rightarrow$ polynomial size circuits of depth $\log^2 n$
(with gates of fan-in 2).

Parallelization algorithm

Let M be the adjcacency matrix of G, add the loop $M_{tt} = 1$.
From undergraduate graphs algorithms:
$\mathrm{output}(G) = (M^p)_{st}$ for any $p \geq \mathrm{depth}(G) = \delta$.
$\Rightarrow$ Compute M^{2^i} for $i = 0, \dots, \log \delta$.

Squaring circuit:
depth 2, m^3 multiplications, m^2 unbounded additions.

General circuits

Theorem[Valiant - Skyum - Berkowitz - Rackoff 1983]:
Let C be a circuit of size s computing a polynomial $f(x_1, ..., x_n)$ of degree d.
There is an equivalent circuit of size $O(d^6 s^3)$ and depth $O(\log(ds) \log d + \log n)$.

Consequence: $\mathsf{VP} \subseteq \mathsf{VNC}^2$ (same as for weakly skew!)

Refinements:

- Uniformity: Miller - Ramachandran - Kaltofen'86; Allender - Mahajan - Jiao - Vinay'98.
- Multilinearity: Raz-Yehudayoff'08.

VP $\subseteq$ VNC^3

The formal degree:

- Multiplication gate: $\deg(f \times g) = \deg(f) + \deg(g)$.
- Addition gate: $\deg(f + g) = \max(\deg(f), \deg(g))$.

Remark:
Formal degree can replace "actual degree" in definition of VP.

Theorem:
Let C be a circuit of size t and formal degree d.
There is an equivalent circuit C' of depth $O(\log t \cdot \log d)$
and size $O(t^3 \log t \cdot \log d)$.
Multiplications gates in C and C' are assumed to be binary.

Remark: if all gates are binary, depth is of order $\log^3$.

Proof of VP $\subseteq$ VNC^3

Let C_i be the "slice" $\{g : \text{gate of } C;\ \deg(g) \in [2^i, 2^{i+1}[\}$.

1. C_i is a (multi-output) circuit with inputs from the C_j $(j < i)$.
2. C_i is skew: if $\deg(g_1), \deg(g_2) \geq 2^i$ then $\deg(g_1 \times g_2) \geq 2^{i+1}$.

Replace each C_i $(i = 0, \ldots, \log d)$
by a circuit of depth $2 \log t$ and size $O(t^3 \log t)$.

Reduction to depth 4 ($\Sigma\Pi\Sigma\Pi$ formulas)

Theorem[Agrawal-Vinay'08]:
Let $P(x_1, \ldots, x_m)$ be a polynomial of degree $d = O(m)$.
If there exists an arithmetic circuit of size $2^{o(d + d \log \frac{m}{d})}$ for P,
then there exists a depth 4 arithmetic circuit of size $2^{o(d + d \log \frac{m}{d})}$.

Corollary:
A multilinear polynomial in m variables with an arithmetic circuit of size $2^{o(m)}$ also has a depth 4 arithmetic circuit of size $2^{o(m)}$.

This suggests to first prove lower bounds for depth 4 circuits.
Warning: For the $n \times n$ permanent, $m = n^2$ and $d = n$.
We already know (Ryser'63) that the permanent
has depth 3 formulas of size $O(n2^n)$!

Reduction to depth 4 for polynomial size circuits

Theorem:
Let C be an arithmetic circuit of size t and formal degree d.
There is an equivalent depth 4 circuit of size $t^{O(\sqrt{d} \log d)}$.

Corollary:
If the permanent family (per_n) is in VP,
then it has depth 4 circuits of size $n^{O(\sqrt{n} \log n)}$.

From branching programs to depth 4 circuits

Theorem:
Let G be an arithmetic branching program of size m and depth δ. There is an equivalent depth 4 circuit with m^2+1 addition gates and $m^{O(\sqrt{\delta})}$ multiplication gates.

Proof: recall $\mathrm{output}(G) = (M^p)_{st}$ for any $p \geq \delta$.

1. Write $M^\delta = (M^{\sqrt{\delta}})^{\sqrt{\delta}}$.
2. Write entries of $N = M^{\sqrt{\delta}}$ as sums of $m^{\sqrt{\delta}-1}$ monomials ($\Rightarrow$ multiplication gates are of arity $\sqrt{\delta}$).
3. Repeat step 2 with matrix M replaced by N.

From general circuits to depth 4 circuits

Start from circuit C of size t and formal degree d, with binary multiplication gates.

1. There is an equivalent branching program G of size $m = t^{\log 2d} + 1$ and depth $\delta = 3d - 1$
2. Convert G into a depth 4 circuit of size $m^{O(\sqrt{\delta})}$.

Proof of step 1:
$C \to$ weakly skew circuit of size $t^{\log 2d}$ (Malod)
$\to$ branching program of size $1 + t^{\log 2d}$;
some additional work for the depth bound.

The τ-Conjecture [Shub-Smale'95]

$\tau(f)$ = length of smallest straight-line program for $f \in \mathbb{Z}[X]$.
No constants are allowed.
Conjecture: f has at most $\tau(f)^c$ integer zeros (for a constant c).
Theorem [Shub-Smale'95]: τ-conjecture $\Rightarrow \mathrm{P}_{\mathbb{C}} \neq \mathrm{NP}_{\mathbb{C}}$.
Theorem [Bürgisser'07]:
τ-conjecture $\Rightarrow$ no polynomial-size arithmetic circuits for the permanent.

Remarks:

- What if constants are allowed?
- We must have $c \geq 2$.
- Conjecture becomes false for real roots: Chebyshev's polynomials, see also Borodin-Cook'76.

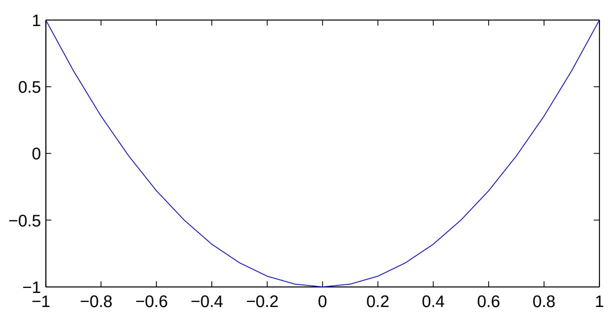
Chebyshev polynomials

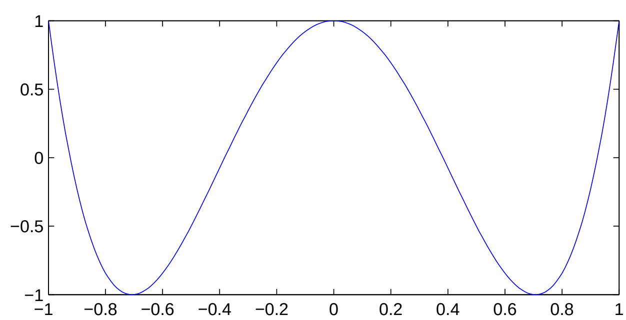
- Let T_n be the Chebyshev polynomial of order n:

$$\cos(n\theta) = T_n(\cos\theta).$$

For instance $T_1(x) = x$, $T_2(x) = 2x^2 - 1$.
- T_n is a degree n polynomial with n real zeros on $[-1, 1]$.
- $T_{2^n}(x) = T_2(T_2(\cdots T_2(T_2(x))\cdots))$: n-th iterate of T_2.
As a result $\tau(T_{2^n}) = O(n)$.

Plots of T_2 and T_4:





The Real τ-Conjecture

Conjecture: Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$,
where the f_{ij} are t-sparse.
If f is nonzero, its number of **real roots** is polynomial in kmt.
Theorem: If the conjecture is true then the permanent is hard.
Remarks:

- It is enough to bound the number of integer roots. Could techniques from real analysis be helpful?
- Case $k = 1$ of the conjecture follows from Descartes' rule.
- By expanding the products, f has at most $2kt^m - 1$ zeros.
- $k = 2$ is open. An even more basic question (courtesy of Arkadev Chattopadhyay): how many real solutions to $fg = 1$? Descartes' bound is $O(t^2)$ but true bound could be $O(t)$.

Descartes's rule without signs

Theorem:
If f has t monomials then f at most $t - 1$ positive real roots.
Proof: Induction on t. No positive root for $t = 1$.
For $t > 1$: let $a_\alpha X^\alpha$ = lowest degree monomial.
We can assume $\alpha = 0$ (divide by X^α if not). Then:

(i) f' has $t - 1$ monomials $\Rightarrow \leq t - 2$ positive real roots.

(ii) There is a positive root of f' between 2 consecutive positive roots of f (Rolle's theorem).

Real τ-Conjecture $\Rightarrow$ Permanent is hard

The 2 main ingredients:

- The Pochhammer-Wilkinson polynomials:
  $PW_n(X) = \prod_{i=1}^{n}(X - i)$.
  Theorem [Bürgisser'07-09]: If the permanent is easy, PW_n has circuits size $(\log n)^{O(1)}$.
- Reduction to depth 4 for arithmetic circuits (Agrawal and Vinay, 2008).

The second ingredient: reduction to depth 4

Depth reduction theorem (Agrawal and Vinay, 2008):
Any multilinear polynomial in n variables with an arithmetic circuit of size $2^{o(n)}$ also has a depth four ($\Sigma\Pi\Sigma\Pi$) circuit of size $2^{o(n)}$.

Our polynomials are far from multilinear, but:

Depth-4 circuit with inputs of the form X^{2^i}, or constants *(Shallow circuit with high-powered inputs)*

$\Updownarrow$

Sum of Products of Sparse Polynomials

How the proof does *not* go

Assume by contradiction that the permanent is easy.
Goal:
Show that SPS polynomials of size $2^{o(n)}$ can compute $\prod_{i=1}^{2^n}(X-i)$
$\Rightarrow$ contradiction with real τ-conjecture.

1. From assumption: $\prod_{i=1}^{2^n}(X-i)$ has circuits of polynomial in n (Bürgisser).
2. Reduction to depth 4 $\Rightarrow$ SPS polynomials of size $2^{o(n)}$.

What's wrong with this argument:
No high-degree analogue of reduction to depth 4
(think of Chebyshev's polynomials).

How the proof goes (more or less)

Assume that the permanent is easy.
Goal:
Show that SPS polynomials of size $2^{o(n)}$ can compute $\prod_{i=1}^{2^n}(X-i)$
$\Rightarrow$ contradiction with real τ-conjecture.

1. From assumption: $\prod_{i=1}^{2^n}(X-i)$ has circuits of polynomial in n (Bürgisser).
2. Reduction to depth 4 $\Rightarrow$ SPS polynomials of size $2^{o(n)}$.

For step 2: need to use again the assumption that perm is easy.

The limited power of powering (a tractable special case)

What if the number of distinct f_{ij} is very small (even constant)?
Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_j^{\alpha_{ij}}(X)$,
where the f_j are t-sparse.

Theorem [with Grenet, Portier and Strozecki]:
If f is nonzero, it has at most $t^{O(m.2^k)}$ real roots.

Remarks:

- For this model we also give a permanent lower bound and a polynomial identity testing algorithm ($f \equiv 0$?). See also [Agrawal-Saha-Saptharishi-Saxena, STOC'2012].
- Bounds from Khovanskii's theory of fewnomials are exponential in k, m, t.

Today's result:

Theorem [with Portier and Tavenas]:
If f is nonzero, it has at most $t^{O(m.k^2)}$ real roots.
The main tool is...

The Wronskian

Definition: Let $f_1, \ldots, f_k : I \to \mathbb{R}$. Their *Wronskian* is the determinant of the *Wronskian matrix*

$$W(f_1, \ldots, f_k) = \det \begin{bmatrix} f_1 & f_2 & \cdots & f_k \\ f_1' & f_2' & \cdots & f_k' \\ \vdots & \vdots & & \vdots \\ f_1^{(k-1)} & f_2^{(k-1)} & \cdots & f_k^{(k-1)} \end{bmatrix}$$

- Linear dependence $\Rightarrow W(f_1, \ldots, f_k) \equiv 0$.
- Converse is not always true (Peano, 1889):
  Let $f_1(x) = x^2$, $f_2(x) = x|x|$. Then
  $$W(f_1, f_2) = \det \begin{bmatrix} x^2 & \mathrm{sign}(x)x^2 \\ 2x & 2\mathrm{sign}(x)x \end{bmatrix} \equiv 0.$$
- Converse *is* true for analytic functions (Bôcher, 1900).

The Wronskian and Real Roots

Upper Bound Theorem: Assume that the k wronskians

$$W(f_1), W(f_1, f_2), W(f_1, f_2, f_3), \ldots, W(f_1, \ldots, f_k)$$

have no zeros on I.
Let $f = a_1 f_1 + \cdots + a_k f_k$ where $a_i \neq 0$ for some i.
Then f has at most $k - 1$ zeros on I, counted with multiplicities.
Remark:
Connections between real roots and the Wronksian were known.
Typical application:
Divide $\mathbb{R}$ into intervals where the k wronskians have no zeros.
Case $k = 2$:

1. If $a_2 = 0$, $f = a_1 f_1$ has no zero on I.
2. If $a_2 \neq 0$, write $f = f_1 g$ where $g = a_1 + a_2 f_2/f_1$.
   $g' = a_2(f_2' f_1 - f_2 f_1')/f_1^2 = a_2 \mathrm{W}(f_1, f_2)/f_1^2$ has no zero $\Rightarrow$
   by Rolle's theorem, g has at most 1 zero, and f too.

Linear Dependence for Analytic Functions (1/3)

Theorem [Bôcher]: If $f_1, \ldots, f_k : I \to \mathbb{R}$ are analytic
and $\mathrm{W}(f_1, \ldots, f_k) \equiv 0$, these functions are linearly dependent.
Proof: By induction on k. Pick $J \subseteq I$ where $f_1 \neq 0$. On J:

$$\begin{aligned} & a_1 f_1 + \cdots + a_k f_k \equiv 0 \\ \Leftrightarrow\quad & a_1 + a_2(f_2/f_1) + \cdots + a_k(f_k/f_1) \equiv 0 \\ \Leftrightarrow\quad & a_2(f_2/f_1)' + \cdots + a_k(f_k/f_1)' \equiv 0. \qquad (*) \end{aligned}$$

(*) follows from induction hypothesis and the recursive formula:

$$\mathrm{W}(f_1, \ldots, f_k) = f_1^k \mathrm{W}((f_2/f_1)', \ldots, (f_k/f_1)').$$

To conclude: for analytic functions,
if $f = a_1 f_1 + \cdots + a_k f_k \equiv 0$ on J, then $f \equiv 0$ on I.

Linear Dependence for Analytic Functions (2/3)

Lemma: $\mathrm{W}(f_1 g, f_2 g, \ldots, f_k g) = g^k \mathrm{W}(f_1, f_2, \ldots, f_k)$.
For instance:

$$\mathrm{W}(f_1 g, f_2 g, f_3 g) = \begin{vmatrix} f_1 g & f_2 g & f_3 g \\ (f_1 g)' & (f_2 g)' & (f_3 g)'' \\ (f_1 g)'' & (f_2 g)'' & (f_3 g)'' \end{vmatrix}$$

$$= g \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' g + f_1 g' & f_2' g + f_2 g' & f_3' g + f_3 g' \\ f_1'' g + 2 f_1' g' + f_1 g'' & f_2'' g + 2 f_2' g' + f_2 g'' & f_3'' g + 2 f_3' g' + f_3 g'' \end{vmatrix}$$

$$= g \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' g & f_2' g & f_3' g \\ f_1'' g + 2 f_1' g' & f_2'' g + 2 f_2' g' & f_3'' g + 2 f_3' g' \end{vmatrix}$$

$$= g^2 \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' g + 2 f_1' g' & f_2'' g + 2 f_2' g' & f_3'' g + 2 f_3' g' \end{vmatrix} = g^3 \mathrm{W}(f_1, f_2, f_3).$$

Linear Dependence for Analytic Functions (3/3): The Recursive Formula for the Wronskian

Proposition [Hesse - Christoffel - Frobenius]:
$\mathrm{W}(f_1, \ldots, f_k) = f_1^k \mathrm{W}((f_2/f_1)', \ldots, (f_k/f_1)')$.
From previous lemma:

$$\mathrm{W}(f_1, f_2, f_3) = f_1^3 \mathrm{W}(1, f_2/f_1, f_3/f_1) = f_1^3 \begin{vmatrix} 1 & f_2/f_1 & f_3/f_1 \\ 0 & (f_2/f_1)' & (f_3/f_1)' \\ 0 & (f_2/f_1)'' & (f_3/f_1)'' \end{vmatrix}$$

Hence

$$\mathrm{W}(f_1, f_2, f_3) = f_1^3 \begin{vmatrix} (f_2/f_1)' & (f_3/f_1)' \\ (f_2/f_1)'' & (f_3/f_1)'' \end{vmatrix} = f_1^3 \mathrm{W}((f_2/f_1)', (f_3/f_1)').$$

Proof of Upper Bound Theorem

Theorem: Assume that the k wronskians

$$W(f_1), W(f_1, f_2), W(f_1, f_2, f_3), \dots, W(f_1, \dots, f_k)$$

have no zeros on I.
Let $f = a_1 f_1 + \cdots + a_k f_k$ where $a_i \neq 0$ for some i.
Then f has at most $k-1$ zeros on I, counted with multiplicities.
Proof: By induction on k.
Assume $k \geq 2$ and $a_2, \dots, a_k$ not all 0.
Write $f = f_1 g$ where $g = a_1 + a_2 f_2/f_1 + \cdots + a_k f_k/f_1$.
To apply induction hypothesis to $g' = a_2(f_2/f_1)' + \cdots + a_k(f_k/f_1)'$:
Note

$$\mathrm{W}((f_2/f_1)', \dots, (f_i/f_1)') = \mathrm{W}(f_1, \dots, f_i)/f_1^i$$

has no zero on I.
Hence g' has at most $k-2$ zeros on I,
g and f at most $k-1$ by Rolle's theorem.

Application: Intersection of a plane curve and a line (1/2)

Theorem (Avendano'09):
Let $g = \sum_{j=1}^{k} a_j x^{\alpha_j} y^{\beta_j}$ and $f(x) = f(x, ax+b)$. Assume $f \not\equiv 0$.
If $b/a > 0$ then f has at most $2k-2$ in each of the 3 intervals
$]-\infty, -b/a[$, $]-b/a, 0[$, $]0, +\infty[$.
Remark: This bound is *provably false* for rational exponents.

Set $a = b = 1$ and $f_j(X) = X^{\alpha_j}(1+X)^{\beta_j}$.
The entries of the wronskians are of the form:

$$f_j^{(i)}(X) = \sum_{t=0}^{i} c_{ijt} X^{\alpha_j - t}(1+X)^{\beta_j - i + t}.$$

Factorizing common factors in rows and columns shows

$$\mathrm{W}(f_1, \dots, f_k) = X^{\sum_j \alpha_j - \binom{k}{2}}(1+X)^{\sum_j \beta_j - \binom{k}{2}} \det M$$

where $\det M$ has degree $\leq \binom{k}{2}$.

Application: Intersection of a plane curve and a line (2/2)

Conclusion:
$f(x) = \sum_{j=1}^{k} a_j x^{\alpha_j}(1+x)^{\beta_j}$ has $O(k^4)$ zeros in $]0,+\infty[$.

Proof:
Assume $W(f_1,\ldots,f_k) \not\equiv 0$ (otherwise, there is a linear dependence).
We have k Wronskians, each with $O(k^2)$ zeros in $]0,+\infty[$.
$\Rightarrow O(k^3)$ intervals containing $\leq k-1$ zeros each.

Remarks:

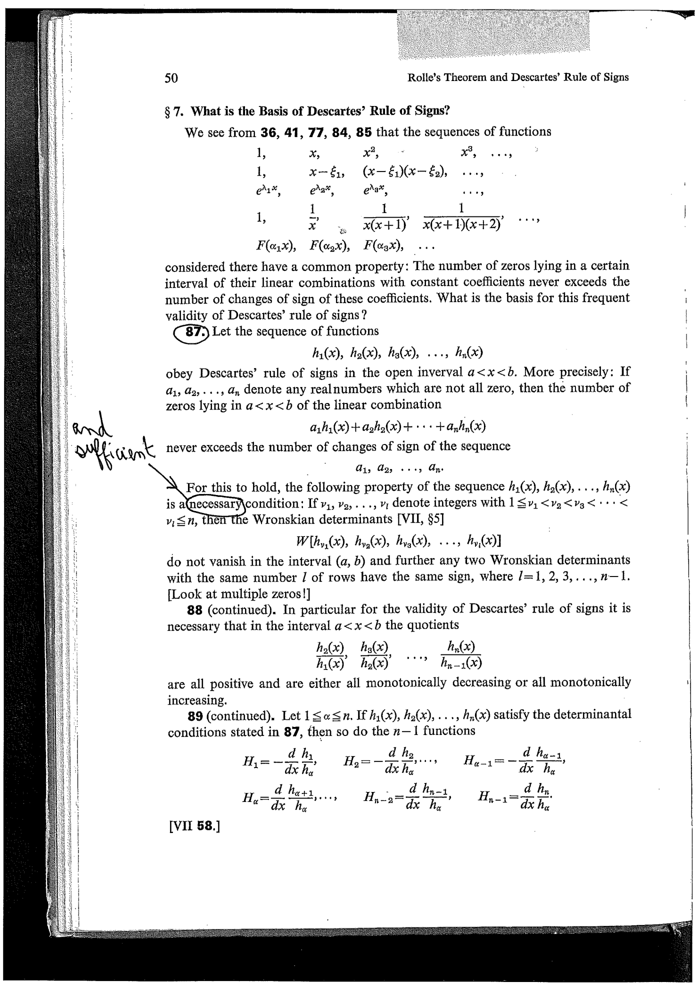
- This can be adapted to a number of different models.
- A better use of the Wronskian leads to $O(k^3)$ upper bound.

To learn more about the Wronskian...

- M. Krusemeyer. Why does the Wronskian work?
  American Math. Monthly, 1988.
  (Recursive formula for the Wronskian)
- A. Bostan and P. Dumas.
  Wronskians and linear independence.
  American Math. Monthly, 2010. *(New non-recursive proof for analytic functions and power series)*
- G. Pólya and G. Szegö.
  Problems and theorems in analysis II.
  (Includes connection to Descartes' rule of signs, pointed out by Saugata Basu)

To learn even more...

- M. Voorhoeve and A. J. van der Poorten.
  Wronskian determinants and the zeros of certain functions.
  Indagationes Mathematicae 78(5):417-424, 1975.
  (*Includes strong version of upper bound theorem; Voorhoeve's papers pointed out by Maurice Rojas*)
- P; Koiran, N. Portier and S. Tavenas.
  A Wronskian approach to the real τ-conjecture.
  `arxiv.org/abs/1205.1015`
  (Preliminary version, check for updates!)



§ 7. What is the Basis of Descartes' Rule of Signs?

We see from **36**, **41**, **77**, **84**, **85** that the sequences of functions

$$1,\quad x,\quad x^2,\quad x^3,\quad \dots,$$
$$1,\quad x-\xi_1,\quad (x-\xi_1)(x-\xi_2),\quad \dots,$$
$$e^{\lambda_1 x},\quad e^{\lambda_2 x},\quad e^{\lambda_3 x},\quad \dots,$$
$$1,\quad \frac{1}{x},\quad \frac{1}{x(x+1)},\quad \frac{1}{x(x+1)(x+2)},\quad \dots,$$
$$F(\alpha_1 x),\quad F(\alpha_2 x),\quad F(\alpha_3 x),\quad \dots$$

considered there have a common property: The number of zeros lying in a certain interval of their linear combinations with constant coefficients never exceeds the number of changes of sign of these coefficients. What is the basis for this frequent validity of Descartes' rule of signs?

87. Let the sequence of functions

$$h_1(x),\ h_2(x),\ h_3(x),\ \dots,\ h_n(x)$$

obey Descartes' rule of signs in the open inverval $a<x<b$. More precisely: If $a_1, a_2, \dots, a_n$ denote any real numbers which are not all zero, then the number of zeros lying in $a<x<b$ of the linear combination

$$a_1h_1(x)+a_2h_2(x)+\cdots+a_nh_n(x)$$

never exceeds the number of changes of sign of the sequence

$$a_1,\ a_2,\ \dots,\ a_n.$$

and sufficient

For this to hold, the following property of the sequence $h_1(x), h_2(x), \dots, h_n(x)$ is a necessary condition: If $\nu_1, \nu_2, \dots, \nu_l$ denote integers with $1 \leqq \nu_1 < \nu_2 < \nu_3 < \cdots < \nu_l \leqq n$, then the Wronskian determinants [VII, §5]

$$W[h_{\nu_1}(x),\ h_{\nu_2}(x),\ h_{\nu_3}(x),\ \dots,\ h_{\nu_l}(x)]$$

do not vanish in the interval (a, b) and further any two Wronskian determinants with the same number l of rows have the same sign, where $l=1, 2, 3, \dots, n-1$. [Look at multiple zeros!]

88 (continued). In particular for the validity of Descartes' rule of signs it is necessary that in the interval $a<x<b$ the quotients

$$\frac{h_2(x)}{h_1(x)},\quad \frac{h_3(x)}{h_2(x)},\quad \dots,\quad \frac{h_n(x)}{h_{n-1}(x)}$$

are all positive and are either all monotonically decreasing or all monotonically increasing.

89 (continued). Let $1 \leqq \alpha \leqq n$. If $h_1(x), h_2(x), \dots, h_n(x)$ satisfy the determinantal conditions stated in **87**, then so do the $n-1$ functions

$$H_1=-\frac{d}{dx}\frac{h_1}{h_\alpha},\quad H_2=-\frac{d}{dx}\frac{h_2}{h_\alpha},\dots,\quad H_{\alpha-1}=-\frac{d}{dx}\frac{h_{\alpha-1}}{h_\alpha},$$
$$H_\alpha=\frac{d}{dx}\frac{h_{\alpha+1}}{h_\alpha},\dots,\quad H_{n-2}=\frac{d}{dx}\frac{h_{n-1}}{h_\alpha},\quad H_{n-1}=\frac{d}{dx}\frac{h_n}{h_\alpha}.$$

[VII **58**.]

Appendix: lower bound for restricted depth 4 circuits

Consider representations of the permanent of the form:

$$\mathrm{per}(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_j^{\alpha_{ij}}(X) \tag{1}$$

where

- X is a $n \times n$ matrix of indeterminates.
- k and m are bounded, and the α_{ij} are of polynomial bit size.
- The f_j are polynomials in n^2 variables, with at most t monomials.

Theorem [with Grenet, Portier and Strozecki]:
No such representation if t is polynomially bounded in n.
Remark: The point is that the α_{ij} may be nonconstant.
Otherwise, the number of monomials in (1) is polynomial in t.

Lower Bound Proof

- Assume otherwise:

$$\mathrm{per}(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_j^{\alpha_{ij}}(X). \tag{2}$$

- Since per is easy, $P_n = \prod_{i=1}^{2^n}(x-i)$ is easy too.
  In fact [Bürgisser], $P_n(x) = \mathrm{per}(X)$ where X is of size $n^{O(1)}$, with entries that are constants or powers of x.
- By (2) and upper bound theorem, P_n should have only $n^{O(1)}$ real roots.
  But P_n has 2^n integer roots!

Remark:
The current proof requires the Generalized Riemann Hypothesis (to handle arbitrary complex coefficients in the f_j).