Polynomial Identity Testing, Michael Forbes

Outline:
1) Overview
   a) algebra
   b) linear maps
   c) sum of powers of linear polynomials
      - lower bounds
      - identity testing
2) sums of powers of quadratic polynomials
   - lower bounds
   - identity testing

1) Overview:

Q (polynomial identity testing (PIT)) given a polynomial \( f(x_1, \ldots, x_n) \), is \( f \equiv 0 \)?

\[ f(x_1, \ldots, x_n) = \sum_{e \in E} c_e x^e \]

where \( x^e \) is a monomial.

For \( g = f \) we can give a polynomial \( h(x) \) by an algebraic circuit.

\[ h(x) = (x-y)(x-y) \]

(Lawrence-Zippel): \( S \in \mathbb{F}_p^x \), \( p \mid f(0) \neq 0 \)

Randomized algo for PIT: given circuit \( C \) compute \( f(0) \)

Use \( C \) to evaluate \( h(x) \) at \( S \), \( |S| = \frac{1}{32} \).

Typical PIT:
- black-box: only use circuit to evaluate \( f(x) \)
- white-box: can use circuit in any way

Long-term goal: efficient deterministic PIT. For interesting models of restricted circuits

- compression: \( f \in \mathbb{F}_p[X] \), \( f \) is \( 20 \) or \( 40 \alg \) and \( f \) computed by small circuits

Conclusion: black box

for circuits of size \( S \), exists \( S \) with \( (S) \leq \text{poly}(S) \)

construct explicit \( S \)?
Q: construct any hitting set?

Lem: \( S \subseteq F \), \( |S| \leq d n \). Then \( S \) is a \((d+1)^n-1\) size hitting set for \( n\)-var, degree \( d \) polynomials. For such polynomials computed by small circuits.

Pf: \( f(x) \) in \( S \), degree \( d \)

\[
S = \{ P_i \mid f(P_i) = 0, 1 \leq i \leq d+1 \}
\]

Hence, \( f(x) \neq 0 \).

Problem: we have small hitting sets if \( n \) is small.

2) Hitting sets \(
\text{fewer} \) than \( d \)-size polynomials,

Prop: \( \exists S \subseteq F \) of \( n\)-var, degree \( d \)

Let \( H \) be a hitting set for \( S \)

Let \( \|H\| < (n+1)^d \) then can find an explicit \( g(x) \) of degree \( d \) in \( \text{poly}(S, H, n, d) \) time.

Given \( |H| \leq \left( \binom{n+d}{d} \right) \) is not hard.

Ideally \( |H| = \text{poly}(S) \)

Hence, \( g(x) \) needs \( \text{size} \approx S \) circuits. Hard to prove.

Pf: \( H \subseteq S \Rightarrow f(x) = 0 \Rightarrow f(x) = 0 \forall x \in S \)

\( f(x) = \sum_{a} c_a x^a \Rightarrow c_a x^a = 0 \forall x \in S \)

\( c_a = 0 \forall a \)

If \( c_a \) is non-zero, linear equation in \( f(x) \).

\( \|H\| \) constraints \( \left( \binom{n+d}{d} \right) \) constraints.

Linear system, so can solve.
Circle at \( \Sigma \) of general circuits.

\[ f(x) = \sum_{i=1}^{n} \left( \alpha_i x_i + \alpha_i x_j + \cdots + \alpha_i x_n \right) \delta_i, \quad \delta_i \in \{0, 1\} \]

"size": \( \text{not a linear polynomial} \)

\[ (x+y)^2 - (x-y)^2 = 4xy \]

---

3) Sum of powers of linear polynomials

defn: a sum of powers of linear polynomials \( \Sigma_{i=1}^{n} \) is

\[ \sum \text{general powering circuits} \equiv \text{general circuit restriction to depth 3 is quite limiting but nontrivial!} \]

\[ \sum x_i = \frac{1}{n!} \sum_{S \subseteq [n]} \left( \sum_{i \in S} x_i \right)^n \]

\[ \text{size} \ n \cdot 2^n \cdot \Sigma \]

**Cor:** any \( n \)-var degree poly has an \( \exp(n^d) \)-size \( \Sigma \) circuit

**Q:** explicit poly nearly large \( \Sigma \) circuits?
3a) \( \Sigma \Lambda \Sigma \) lbs

thing: \( x_1 \cdots x_n \) need \( \Sigma \Lambda (x) \) size \( \Sigma \Lambda \) circuit

strategy: complexity measure \( m: \mathbb{R}[x] \to \mathbb{N} \)

\( s.t.: \) \( m(x_1 \cdots x_n) = \text{large} \)

\( m(f) = \text{small} \) if \( f \) has small \( \Sigma \Lambda \) circuit

\( i.e.: \) \( m(f \circ g) = m(f) + m(g) \)

\( m(\ell^d) = \text{small} \) if \( \deg \ell \leq d \)

\( \Rightarrow \) \( x_1, x_2, \ldots, x_n \in \Sigma \Lambda \) \( \ell \)

then: \( r = \frac{m(x_1 \cdots x_n)}{m(\ell^d)} \Rightarrow \frac{\text{large}}{\text{small}} = \text{large} \)

fundamental measure: "size" of partial derivative

\( \exists x: \) \( f(x_1, \ldots, x_n) \)

\( \Rightarrow \) there is only 1 partial derivative of each order!

\( \partial x_i f = \frac{\partial f}{\partial x_i} (x_1, \ldots, x_n) \)

\( \partial x_i y = \frac{\partial y}{\partial x_i} (x_1, \ldots, x_n) \)

ex: \( f(x, y) = (x + y)^n \)

chain rule: \( \partial x f = (x + y)^{n-1} \cdot 1 \)

\( \partial x y = n(x + y)^{n-1} \cdot 1 \)

\( \partial y = 2x + y \leq 2 \) polynomials essentially the same

idea: measure size by linear algebra

\( C: \mathbb{Q}[x]^{ed} \to (\mathbb{Q}[x])^{ed} \) is a \( \text{dim} \mathbb{Q}[x] = \text{dim} \mathbb{Q}[x] \) linear map

\( \text{def.} (\text{partial derivative method}): \) \( \Delta: \mathbb{Q}[x]^{ed} \to \mathbb{N} \mathbb{Q}[x]^{ed} \)

\( \Delta(f) = \sum \Delta x_i f \)

\( \Delta x_i f = \frac{\partial f}{\partial x_i} (x_1, \ldots, x_n) \)

\( \text{dim} \Delta(f) = \text{dim} \mathbb{Q}[x]^{ed} \)

ex: \( f(x, y) = (x + y)^n \)

\( \partial f = \sum (x + y)^{n-1} \cdot 1 \cdot (x + y)^{n-1} \cdot 1 \cdot \Delta x + \Delta y \)

\( \text{dim} = 3 = 2 + 1 \)

ex: \( f(x, y) = xy \)

\( \partial f = \Delta x, \Delta y, z \)

\( \text{dim} = 4 = 2^2 \)

\( f \) is 3 not 5 or 6...
\[
\begin{align*}
\deg x + \deg y &= \deg (x + y) + \deg (x) + \deg (y), \\
\text{Proof:} \quad \deg x + \deg y &= \deg (x + y) + \deg (x) + \deg (y), \\
\text{Thus} \quad \deg (x + y) &\in \text{Span}\{x, y\}.
\end{align*}
\]

Lemma: \( x + y \in \text{Span}\{x, y\} \)

Proof: \( x + y = \sum a_i x_i + \sum b_i y_i \)

So \( x + y \in \text{Span}\{x, y\} \)

Lemma: \( x^2 + y^2 \in \text{Span}\{x, y\} \)

Proof: \( x^2 + y^2 = (x + y)^2 - 2xy \)

So \( x^2 + y^2 \in \text{Span}\{x, y\} \)

This "clones" the lower bounds question but not identically steering: no hardness as answers are explainable.

Theorem: given \( \deg \) size \( \Sigma \Lambda \), can decide whether \( f \equiv 0 \) in \( \text{poly}(\deg) \).

Proof: there are \( \text{poly}(\deg) \) \( \deg \)-size hitting sets for \( \Sigma \Lambda \) circuits.

Theorem: \( \text{poly}(\deg) \) \( \deg \) cross-bounds despite simplicity of \( \Sigma \Lambda \).

Here: "scale down" with low bandwidth "see more" from lower bound.
$$f(x) = \sum_{i=1}^{d} f_i(x), \quad d \leq d_{\text{col}}, \quad \text{what does this mean?}$$

If $1 < s^d$ then $f \not\propto x_1 \cdots x_n$

Sub-case:
$$f \propto x_1 \cdots x_n, \quad i_1 < \cdots < i_k \quad \text{to } k \gg \log v$$

So $\sum \langle \text{size of } f \rangle$ can only compute $\log(v)$-size monomials.

Let
$$(x_1 + \cdots + x_n)^n = x_1^n + x_2^n + \cdots + x_n^n$$

Let $\sigma$ be a large random $\mathbb{F}_p$-only vector.

When $f = \sum_{\gamma \in \mathbb{F}_p^d} \mathbb{F}_p \gamma$,

Then $\langle \text{size of } f \rangle \leq \log(n)$,

$$\text{since } \sum \langle \text{size of } f \rangle \text{ involves } \leq \log(n) \text{ variables}$$

In this, $\varnothing$ is the list of non-zero monomials.

The lexicographic ordering on monomials $x$ is \[ \text{dictionary order} \]

Let $f = \sum_{\gamma \in \mathbb{F}_p^d} \mathbb{F}_p \gamma$

Let $x > y > 2^{100}$

$$x^2 + y^2 > 2^{200}$$

Thus, $\langle x_1 + \cdots + x_n \rangle = x_1^n$ for some order $+ \cdots + x_n^n$

This explains how one can $f$ in size $\sum \langle f \rangle$

$P^h = 1)$ guess the $\langle \text{size of } f \rangle$

2) zero out other variables with lower monomial.

3) get $\varnothing$ by poly in $\langle \text{size of } f \rangle$ vars

4) build true hitting set (random poly by $\langle \text{size of } f \rangle$ choices.

Formally, $S \subseteq \mathbb{F}_p^d$, $S = \langle \text{size of } f \rangle + \langle \text{size of } f \rangle$, $S$ is hereditary, closed, poly

$$\varnothing = \big\{ \mathbb{F}_p^d \big\}, \quad \langle \text{size of } f \rangle \text{ hits } \varnothing$$

Let $\varnothing$ is a total order on monomials.

1) \[ x^d < x^d \]

2) \[ x^a < x^b \text{ if } \langle a \rangle < \langle b \rangle \]

3) \[ x^a < x^b \iff h(x^a) < h(x^b) \]

$P^h = 1)$ by olth

2) \[ \langle a \rangle < \langle b \rangle \quad \text{for nonzero words} \]

3) \[ \langle \text{first nonzero word of } f \rangle \text{ is positive, e.g. } f(x_1, x_2) \]
Define $x^2 < x^3$. Suppose $\frac{\partial^e(x^2)}{\partial x^e(x^2)} \neq 0$.

**Claim:** $\frac{\partial^e(x^2)}{\partial x^e(x^2)} < \frac{\partial^e(x^3)}{\partial x^e(x^3)}$.

**Proof:**
\[
\frac{\partial^e(x^2)}{\partial x^e(x^2)} = \text{const.}, \quad \frac{\partial^e(x^3)}{\partial x^e(x^3)} = \frac{\partial^e(x^2)}{\partial x^e(x^2)} + \text{lower term}
\]
This implies $x^2 < x^3$.

**Main Lemma:** \( \left| \frac{\partial^e(f)}{\partial x^e(LM(f))} \right| \)

When \( LM(f) \) is the leading monomial of \( f \),

so if \( x = x^2 \text{ then } f = (x, x^2, \text{ lower terms}) \)

**Main Lemma:** Suppose \( f \) is a polynomial.

Let \( x, y, z \in \mathbb{R} \) such that \( \frac{\partial^e(f)}{\partial x^e(LM(f))} \geq 2^k \)

If \( LM(f) \) involves \( x \) only \( = \) leading monomial involves \( x \) or lower terms.

**Proof of Main Lemma:**

Let \( f \) be a polynomial.

\( \frac{\partial^e(f)}{\partial x^e(LM(f))} \geq 2^k \) if \( x^k \) in \( LM(f) \).

**Claim:**
\( f = x^k + \text{lower} \)

\( \frac{\partial^e(f)}{\partial x^e(LM(f))} = 0 \) if \( \frac{\partial^e(x^k)}{\partial x^e(x^k)} \geq 0 \).

**Claim:** All polynomials in \( f \) are linearly independent.

**Proof:** \( LM(\frac{\partial^e(f)}{\partial x^e(LM(f))}) \geq 2^k \) such that \( x^k \) involves \( x \).
Recall: defined \( \Sigma \Lambda \Sigma \) - \( x_1, \ldots, x_n \) has size \( 2^n \Sigma \Lambda \Sigma \\
\) - define \( |\Sigma| \) to be complexity measure \\
\( |\Sigma(x_1, \ldots, x_n)| = 2^n \)
\( = \) read \( 2^n \) size \( \Sigma \Lambda \Sigma \) - \( x_1, \ldots, x_n \)
\( \) (sizes \( \Sigma \Lambda \Sigma \)) \( \leq S \)

- "Scale down" and "go slow" from lower bound
- if sizes \( \Sigma \Lambda \Sigma \) = \( \text{LM}(P) \) involves \( O(g(r)) \) (arithmetic)

- use structural results to gain binding size

2) sums of powers of quad polynomials

be studied \( \Sigma \Lambda \Sigma^2 \) when \( \deg(x_i) \leq 1 \) \( (\Sigma \Lambda \Sigma) \)

Q: when if \( \deg(x_i) > 2 \) \( \in \Sigma \Lambda \Sigma \Sigma^2 \)

Scenarios:

1) \( \Sigma \Lambda \Sigma \Sigma^2 \) size \( S = \Sigma \Lambda \Sigma \) grew \( \leq g(r) \)

2) some \( f \) small \( \Sigma \Lambda \Sigma \Sigma^2 \) - \( f \) needs large \( \Sigma \Lambda \Sigma \)

[Diagram:
\( \Sigma \Lambda \Sigma \) scenario: \( X_1, \ldots, X_n \) has small \( \Sigma \Lambda \Sigma \Sigma^2 \) circuit
\( \Rightarrow X_1, \ldots, X_n \) needs large \( \Sigma \Lambda \Sigma \Sigma^2 \) growth over \( 2^{\alpha \cdot r} \).

Prep: \( f = (x_1^2 + \ldots + x_n^2)^n \)

a) \( f \) has small \( \Sigma \Lambda \Sigma \Sigma^2 \)

b) \( f \) needs \( 2^{\alpha \cdot r(n)} \Sigma \Lambda \Sigma \)

Pf: (c) by direct

(c) \( \max \{|\Sigma|\} = 2^n \)

\( f(x_1, y) = (x^2 + y^2)^n \)

\( \partial_x = 2x(x^2 + y^2) \quad x \)

\( \partial_y = 2y(x^2 + y^2) \quad y \)

\( f_{x,y} = x^2y \quad xy \)

Claim: \( S = 2^n \) \( \prod \pi \in C \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \) constant. \( \prod \pi \in \Sigma \Sigma \Lambda \Sigma \Sigma^2 \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \)

Pf: induction:

\( S = \#S \)

\( S = \sum (x_1^2 + \ldots + x_n^2)^n \)

\( S = \sum \prod \pi \in C \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \)

\( \Rightarrow S = \sum \prod \pi \in C \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \lambda \Sigma \Sigma \Lambda \Sigma \Sigma^2 \)

\( = 2(x_1^2 + \ldots + x_n^2)^n \)
Polynomial Identity Testing, Michael Forbes

Outline:
1) Review
   - Identity testing of sums of powers of linear polys
   - Sums of powers of quadratic polys (?)
2) Review

Q: polynomial identity testing: given a polynomial $f(x)$, is $f(x^n)$ polynomially identity testing?
Theorem: $x_1 - x_n$ needs $2 \log(n)$ size $\mathbb{Z}^n$.

Thus: the leading moments of size $s \mathbb{Z}^n$ yields $s \ll n^{1/2}$.

Proof: use $z \in \mathbb{Z}^n$ to get $y_1, \ldots, y_n$.

Then substitute $y_1 = x_1 - x_d$,

$y_2 = x_d - x_1$,

$\vdots$

$y_n = x_n - x_1$,

Thus: $x_1 - x_n$ needs $2 \log(n)$ size $\mathbb{Z}^n$.

Only in 2012!

Thus: the leading moments of size $s \mathbb{Z}^n$ yields $s \ll n^{1/2}$.

Claim: the leading moments of size $s \mathbb{Z}^n$ yields $s \ll n^{1/2}$.

Insight: consider $Q_{x_1, \ldots, x_n}^\infty = \sum_{i=1}^n P_{x_i} = \sum_{i=1}^n P(x_1, x_2, \ldots, x_n)$.

Both sets are of same size but (a) is easy to use and (b) is easy to use.

Key property of $\deg s$: low $\deg s$, low $\deg s$.

New measure: $x = \deg s(f) = \begin{cases} x_0^a & x_0 \leq a \\ \deg x^s \leq x \end{cases}$.

"Shifted partial orderings."

Lots of exciting recent work on this measure.

Proving the 2nd is not too hard, but another