Lower bound for circuits with MOD$_p$ gate

- [Razborov '87; Smolensky '87]

$$MOD_p(x_1, x_2, \ldots, x_n) = \begin{cases} 1 \text{ if } \sum_{i=1}^{n} x_n \neq 0 \pmod p \\ 0 \qquad \text{otherwise} \end{cases}$$

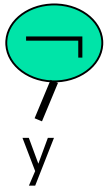
- *Result* :

  p, q: powers of distinct primes
  the MOD_p function cannot be computed with $\{AND, OR, NOT, MOD_q\}$ of poly size and constant depth

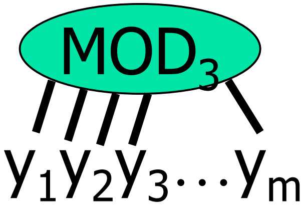
- $Theorem:$ (a special case with $p = 2$ and $q = 3$) MOD_2 function cannot be computed with a circuit of size $\frac{1}{10}2^{.5n^{\frac{1}{2d}}}$, depth d by using AND, OR, NOT and MOD_3-gates for sufficiently large n

$\swarrow \{\wedge, \vee, \neg, MOD_3\}$

- Let C be a circuit of depth d computing MOD_2 Approximate C with small degree poly in $GF(3)$ by replacing each gate in C with an approximate polynomial

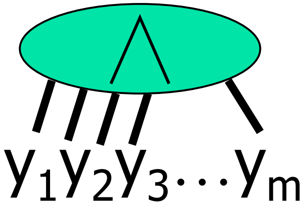
- 

$\Rightarrow (1-y)$

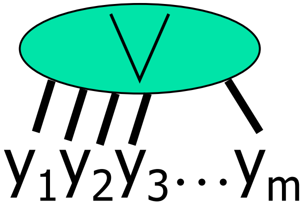
- 

$\Rightarrow (y_1+\ldots+y_m)^2$

No error introduced

- 

$\Rightarrow y_1 y_2 \ldots y_m$

- 

$$\Rightarrow \overline{\left(\bar{y}_1 \wedge \ldots \wedge \bar{y}_m\right)} = 1-(1-y_1)\ldots(1-y_m)$$

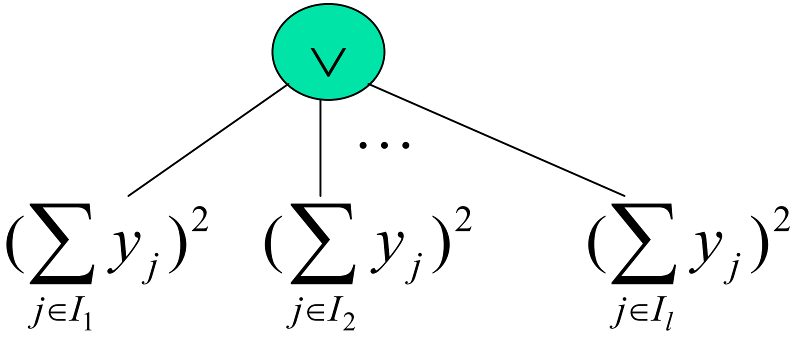
degree increases a lot

- l : an integer to be determined later
Choose l random subsets $I_1,\ldots,I_l \subseteq \{1,2,\ldots,m\}$
with $\Pr[i \in I_j] = \frac{1}{2}$

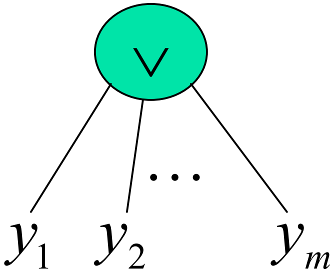
- If $OR(y_1, y_2, \ldots, y_m) = 0 \quad \Rightarrow \forall 1 \le i \le l \quad (\sum_{j \in I_i} y_j)^2 = 0$

  If $OR(y_1, y_2, \ldots, y_m) = 1 \quad \Rightarrow \forall i \quad \Pr[(\sum_{j \in I_i} y_j)^2 = 1] \ge \frac{1}{2}$

- $\frac{1}{2^m}(\binom{m'}{0} + \binom{m'}{3} + \ldots)2^{m-m'}$

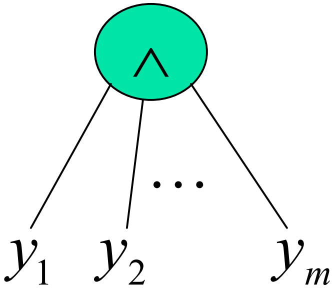

$\vee$
$\cdots$
$(\sum_{j \in I_1} y_j)^2$
$(\sum_{j \in I_2} y_j)^2$
$(\sum_{j \in I_l} y_j)^2$

- If $OR(y_1, y_2, \dots, y_m) = 0$ then the above output 0
  If $OR(y_1, y_2, \dots, y_m) = 1$ then $\Pr[output\ 1] \geq 1 - 2^{-l}$



$$\Rightarrow 1 - \prod_{i=1}^{l} (1 - (\sum_{j \in I_i} y_j)^2)$$

$\deg = 2l$, but introduce some error



$$\Rightarrow \prod_{i=1}^{l} (1 - [\sum_{j \in I_i} (1 - y_j)]^2)$$

by De Morgan's rules

$$y_1 \wedge \dots \wedge y_m = \overline{\overline{y_1} \vee \dots \vee \overline{y_m}}$$

- For circuit C of depth d and size s, we obtain a polynomial P of degree at most $(2l)^d$

- $\Pr[$each poly agree with the corresponding gate of C $] \geq 1 - \frac{s}{2^l}$

- $Exp(|\{(x_1,\ldots,x_n) : P(x_1,\ldots,x_n) = C(x_1,\ldots,x_n)\}|) \geq 2^n(1 - \frac{s}{2^l})$

- *Lemma* 1:
  C : as defined above
  ℓ : any positive integer
  $\exists$ a polynomial P over $GF(3)$ has deg $\le (2\ell)^d$ and

  $$\left|\{(x_1,\ldots,x_n) : P(x_1,\ldots,x_n) = C(x_1,\ldots,x_n)\}\right| \ge 2^n(1-\frac{s}{2^\ell})$$

- *Lemma 2:*
  There is no polynomial $P(x_1,\ldots,x_n)$ over $GF(3)$ of degree at most $\sqrt{n}$ that is equal to the parity of $x_1,\ldots,x_n$ for a set S of at least $0.9\cdot 2^n$ distinct binary vectors $(x_1,\ldots,x_n)$

- *proof:*
  By contradiction, suppose $\exists$ a P over $GF(3)$, s.t.

  $P(x_1,\ldots,x_n)=x_1\oplus\ldots\oplus x_n$ for all $(x_1,\ldots,x_n)\in S$

  $\deg(P)\le\sqrt{n}$

  Define $Q(y_1,\ldots,y_n)=P(y_1+2,\ldots,y_n+2)-2$

  So $P=1\rightarrow Q=-1;\quad P=0\rightarrow Q=1(\equiv -2 \mod 3)$

  $T=\{(y_1,\ldots,y_n)\in\{1,-1\}^n:(y_1+2,\ldots,y_n+2)\in S\}$

  Thus $\deg(Q)\le\sqrt{n},\quad Q(y_1,\ldots,y_n)=\prod_{i=1}^{n}y_i \; over \; T$

- Consider an arbitrary function $G(y_1,\dots,y_n):T\rightarrow GF(3)$, and extend it to a function from $(GF(3))^n\rightarrow GF(3)$, which is a polynomial.
- Replace each y_i^2 with 1 and obtain $\widetilde{G}$, which agree with G on T.
- Replace each $\prod_{i\in U} y_i$, where $|U|>\frac{n}{2}+\frac{\sqrt{n}}{2}$ by $\prod_{i\notin U} y_i\cdot Q(y_1,y_2,\dots,y_n)$ and remove y_i^2 to obtain $\widetilde{\widetilde{G}}$ which is equal to G on T and $\deg(\widetilde{\widetilde{G}})\le\frac{n}{2}+\frac{\sqrt{n}}{2}$

$$\because \prod_{i\notin U} y_i\cdot\prod_{i=1}^{n} y_i=\prod_{i\in U} y_i$$

# of possible $\tilde{\tilde{G}}$:

- $$\underbrace{a_0}_{\binom{n}{0}} + \underbrace{a_1x_1 + \cdots + a_nx_n}_{\binom{n}{1}} + \underbrace{\sum a_{ij}x_ix_j}_{\binom{n}{2}} + \cdots + \underbrace{\sum a_{i_1\ldots i_n}x_{i_1}\cdots x_{i_{\frac{n}{2}+\frac{\sqrt{n}}{2}}}}_{\binom{n}{\frac{n}{2}+\frac{\sqrt{n}}{2}}}$$

$$3^{\sum_{i=0}^{\frac{n}{2}+\frac{\sqrt{n}}{2}}\binom{n}{i}} < 3^{.8\cdot 2^n} \qquad why?$$

$$G \to \tilde{G} \to \tilde{\tilde{G}}$$

- But # of possible $G: \ T \to GF(3)$:

$$3^{|T|} > 3^{.9\cdot 2^n} \qquad \to\leftarrow$$

■

- *Corollary* :

  No circuit of depth d and size s $\le \frac{1}{10} 2^{0.5 n^{\frac{1}{2d}}}$ compute $x_1 \oplus x_2 \oplus \cdots \oplus x_n$ using NOT, $\wedge$, $\vee$, MOD_3-gates

- *proof* :

  Suppose not, let C be such a circuit. Let $l = \frac{1}{2} n^{\frac{1}{2d}}$
  By Lemma 3.1, $\exists$ a poly. $P(x_1,\ldots,x_n)$ over $GF(3)$ whose degree $\le (2d)^d = \sqrt{n}$, which is equal to the parity of $x_1,\ldots,x_n$ on at least $2^n(1-\frac{S}{2^{n^{\frac{1}{2d}}/2}}) \ge .9 \cdot 2^n$ inputs

  Contradicts Lemma 2

■

- $Open\ \ question:$
  - $*$ What is the lower bound of computing MOD_5 with $\{\wedge, \vee, \neg, MOD_6\}$?
  - $*$ What is the computing power of MOD_m-gate, when m is not prime or prime power?

- $Note:$

  Properties of finite fields do not work in this, since Z_m is not a field when m is not a prime power!

- Monotone formulas for all threshold functions:

  For the majority function

$$f(x_1,\ldots,x_n)=\begin{cases}1, & \text{if } \geq \frac{n}{2} \text{ of } x_i\text{'s are } 1\\ 0, & \text{otherwise}\end{cases}$$

Formula size:

$$\begin{cases}\Omega(n^2), \text{ best known over } \{\wedge, \vee\}, \{\wedge, \vee, \neg\}\\ O(n^{5.271}), \text{ over } \{\wedge, \vee\}, \text{ by Valiant, 84,}\\ \qquad\qquad\qquad\qquad \text{in J. of algorithm}\end{cases}$$

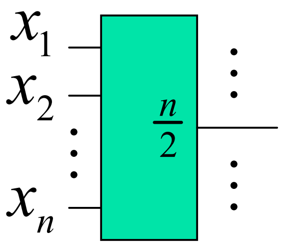
Circuit size:

sorting network: [AKS] (Ajtai, Komlos, Szemeredi)'83 STOC

Depth: $O(\log n)$

Size: $O(n \log n)$

n: huge constant



- *Theorem* :
  The monotone formula size of each threshold function T_k^n is $O(n^{5.271})$

$$T_k^n(x_1, x_2, ..., x_n) = \begin{cases} 1, \text{ if } x_1 + ... + x_n \geq k \\ 0, \text{ otherwise} \end{cases}$$

Construct random formulas level by level:

Level 0: For some $0 < p < 1$
choose x_j $(1 \leq j \leq n)$ as a formula with probability p; 0 with probability $1 - np$

Level i-1: G_i's ...

Level i : $F_i = (G_1 \vee G_2) \wedge (G_3 \vee G_4)$
independently chosen from level i-1

$size(F_i) \leq 4^i$

- If $\Pr[F_i = T_m^n] > 0$, then $\exists$ a monotone formula of $size \le 4^i$ for T_m^n

- It is sufficient to prove that

$$\forall \vec{a} \in \{0,1\}^n, \Pr[F_i(\vec{a}) \neq T_m^n(\vec{a})] < 2^{-n-1}$$

- $\Pr[\exists \vec{a}, \quad F_i(\vec{a}) \neq T_m^n(\vec{a})]$
$\le \sum_{\vec{a} \in \{0,1\}^n} \Pr[F_i(\vec{a}) \neq T_m^n(\vec{a})] \le 2^n \bullet 2^{-n-1} = \frac{1}{2}$

- $F_i = (G_1 \vee G_2) \wedge (G_3 \vee G_4)$

$$f_i = \max\{\Pr[F_i(\vec{a}) = 1] : T_m^n(\vec{a}) = 0\} = \max_{\substack{\vec{a} \in \{0,1\}^n \\ T_m^n(\vec{a})=0}} \Pr[F_i(\vec{a}) = 1]$$

$$h_i = \max\{\Pr[F_i(\vec{a}) = 0] : T_m^n(\vec{a}) = 1\} = \max_{T_m^n(\vec{a})=1} \Pr[F_i(\vec{a}) = 0]$$

- *Lemma* :

  $$f_i = f_{i-1}^4 - 4f_{i-1}^3 + 4f_{i-1}^2$$

  $$h_i = -h_{i-1}^4 + 2h_{i-1}^2$$

  proof :

  F_i: monotone, symmetric

  F_i has its worst behavior on inputs with exactly m or m-1 1's

  Let $\vec{a}$ be an input with m-1 1's

  $$T_m^n(\vec{a}) = 0$$

  G_j: $(i$-$1)$-th level formula, for $j = 1,\ 2,\ 3,\ 4$

Thus $f_{i-1} = \Pr[G_j(\vec{a}) = 1], \quad j = 1,\ 2,\ 3,\ 4$

$$\Pr[(G_1(\vec{a}) \vee G_2(\vec{a})) = 1]$$

$$= 1 - (1 - f_{i-1})^2 = -f_{i-1}^2 + 2f_{i-1}$$

$$= \Pr[(G_3(\vec{a}) \vee G_4(\vec{a})) = 1]$$

$$\therefore \quad \Pr[F_i(\vec{a}) = 1] = (-f_{i-1}^2 + 2f_{i-1})^2$$

$$= f_{i-1}^4 - 4f_{i-1}^3 + 4f_{i-1}^2 \quad (< 4f_{i-1}^2)$$

$$h_i = \max\{\Pr[F_i(\vec{a}) = 0] : T_m^n(\vec{a}) = 1\}$$

$$h_i = -h_{i-1}^4 + 2h_{i-1}^2 \qquad (< 4h_{i-1}^2)$$

Let $\vec{b}$ be an input with m 1's

$$T_m^n(\vec{b}) = 1$$

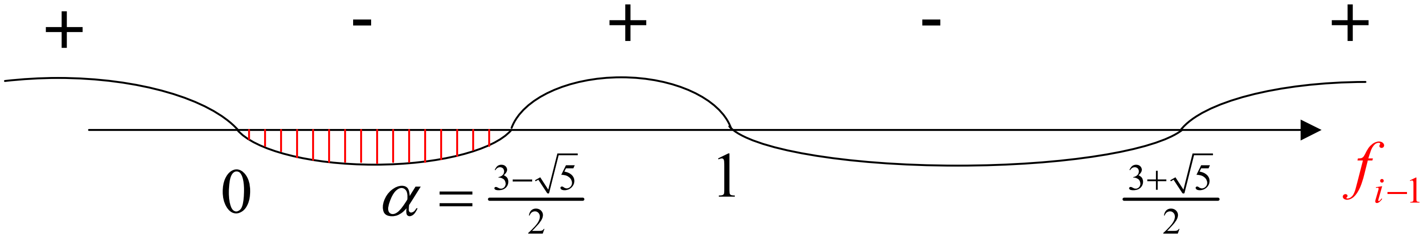
$$h_{i-1} = \Pr[G_j(\vec{b}) = 0], \quad j = 1,\ 2,\ 3,\ 4$$

$$\Pr[(G_1(\vec{b}) \vee G_2(\vec{b})) = 0] = h_{i-1}^2$$

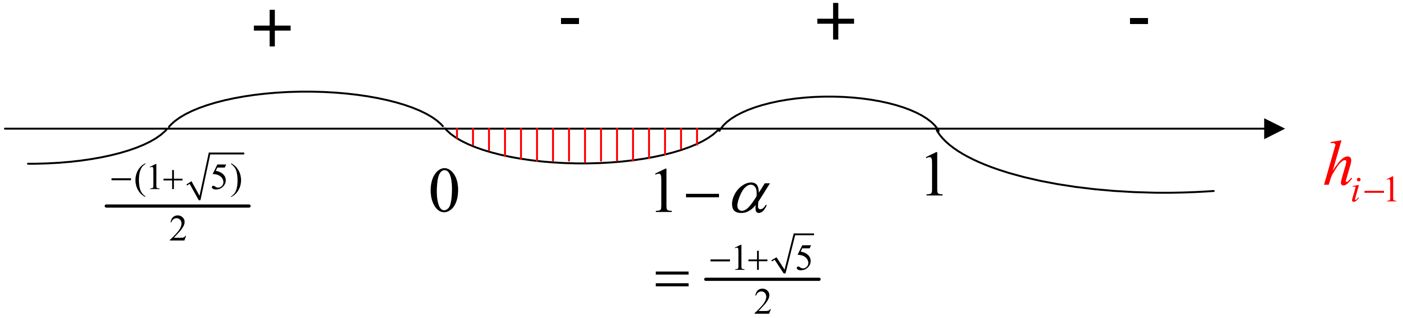
$$\begin{aligned} h_i &= \Pr[(G_1(\vec{b}) \vee G_2(\vec{b})) \wedge (G_3(\vec{b}) \vee G_4(\vec{b})) = 0] \\ &= 1 - (1 - h_{i-1}^2)^2 = -h_{i-1}^4 + 2h_{i-1}^2 \end{aligned}$$

$\blacksquare$

- What f_i's, h_i's decrease!
  - $f_i - f_{i\text{-}1} \ < \ 0,$ i.e. $f_{i-1}^4 - 4f_{i-1}^3 + 4f_{i-1}^2 - f_{i-1} < 0$ $\quad 0 \le f_{i-1} \le 1$



  - $h_i - h_{i\text{-}1} \ < \ 0,$ i.e. $-h_{i-1}^4 + 2h_{i-1}^2 - h_{i-1} < 0$ $\quad 0 \le h_{i-1} \le 1$



- Take $p = 2\alpha/(2m-1)$

  Let $\vec{a}$ have $m-1$ 1's

  $$f_0 = \Pr[F_0(\vec{a}) = 1] = (m-1)p = \alpha - \frac{\alpha}{2m-1} \le \alpha - \Omega(\tfrac{1}{n})$$

  Let $\vec{b}$ be an input with m 1's

  $$\begin{aligned} h_0 = \Pr[F_0(\vec{b}) = 0] &= 1 - \Pr[F_0(\vec{b}) = 1] \\ &= 1 - mp \\ &= 1 - \alpha - \frac{\alpha}{2m-1} \le 1 - \alpha - \Omega(\tfrac{1}{n}) \end{aligned}$$

  We know $f_i < 4f_{i-1}^2$ & $h_i < 4h_{i-1}^2$ from the above lemma

  Thus

  $$f_l < 4f_{l-1}^2 < 4^3 f_{l-2}^4 < \ldots < 4^{2^i-1} f_{l-i}^{2^i} \le 4^{-n-1} < 2^{-n-1}$$

  by taking $f_{l-1} < \frac{1}{16}$ and $n = 2^i$ $(\log n = i)$

- Let δ be a small positive number

Then $f_{i-1} = \alpha - \delta \quad \Rightarrow f_i = \alpha - 4\alpha\delta + O(\delta^2) < \alpha - r\delta$

$\because \; f_0 = \alpha - c\frac{1}{n} \quad \Rightarrow f_j < \alpha - r^j \cdot \frac{c}{n}$

So $f_{l-i} < \alpha - r^{l-i} \cdot \frac{c}{n} = \frac{1}{16}$

$\Rightarrow (\alpha - \frac{1}{16}) = \frac{c}{n} \cdot r^{l-i} \Rightarrow c' + \log n = (l-i) \cdot \log r$

$$\Rightarrow l = i + \frac{\log n}{\log r} + c'' = \log n + \frac{\log n}{\log r} + c'', \qquad \frac{1}{\log r} \approx 1.63$$

$$4^l \approx 4^{\log n + 1.63 \log n + c''} = O(n^{5.26})$$

- $Similarly:$

$$h_l < 4h_{l-1}^2 < \ldots < 4^{2^i-1}\cdot h_{l-i}^{2^i} \le 4^{-n-1} < 2^{-n-1}$$

by taking $h_{l-i} < \frac{1}{16}$ and $n = 2^i$

$$\because\ h_{i-1} = 1-\alpha-\delta \Rightarrow h_i = 1-\alpha-4\alpha\delta+O(\delta^2)$$

$$< 1-\alpha-r\delta,\qquad r<4\alpha$$

And $h_0 \le 1-\alpha-\frac{c}{n}$

$$\Rightarrow h_j < 1-\alpha-r^j\cdot\frac{c}{n}$$

So, $h_{l-i} < 1-\alpha-r^{l-i}\cdot\frac{c}{n} = \frac{1}{16}$

$$\Rightarrow 1-\alpha-\frac{1}{16} = \frac{c}{n}\cdot r^{l-i}$$

$$\Rightarrow c'+\log n = (l-i)\log r \Rightarrow \quad l = i+\frac{\log n}{\log r}+c''$$

$$\approx \log n + 1.63\log n + c''$$

Thus $4^l = O(n^{5.26})$

$$\therefore\ Size(F_l) = O(n^{5.26})$$ ■

- *remarks* :

  The best Lower bound on the formula size of the majority function (for monotone $\&$ $\{\wedge, \vee, \neg\}$ basis) : $\Omega(n^{2}) \leftarrow ---------\rightarrow O(n^{5.2\ldots})$

  Can this construction be applied on the other problems?

- $f_{i-1} \le \alpha - \delta$, where δ is a small positine number
  $f_i = \alpha - 4\alpha\delta + O(\delta^2)$, by the recursive equation
  $\le \alpha - 4\alpha\delta + \delta^2$
  $\le \alpha - \underbrace{(4\alpha - \delta)}_{r_i}\delta$

  if is every small, then $r_i \approx 1.\cdots$, since $4\alpha = 1.\cdots$

  $$f_{i-1} \le \alpha - 4\alpha r_i\delta + r_i^2\delta^2$$
  $$= \alpha - r_i\underbrace{(4\alpha - r_i\delta)}_{r_{i+1}}\delta$$
  $$= \alpha - r_i r_{i+1}\delta, \qquad r_i > r_{i+1} < 4\alpha$$
  $$< \alpha - r_{i+1}^2\delta$$