Foundations of Probabilistic Proofs

A course by **Alessandro Chiesa**

Lecture 02

Sumcheck Protocol

Interactive Proofs for Counting Problems

We saw an IP for GNI, a problem in coNP not known to be in P.

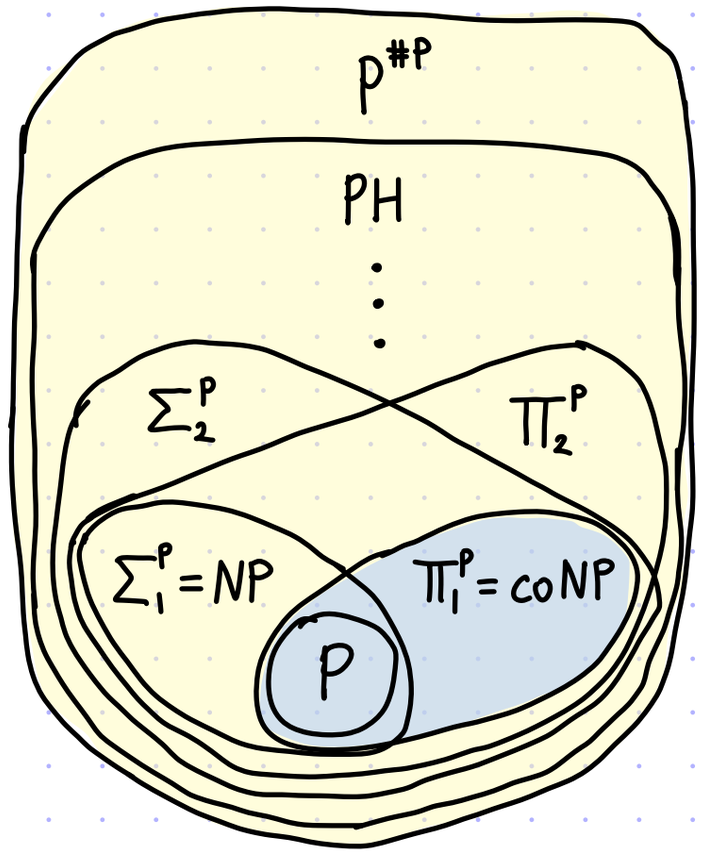
But GNI is not believed to be coNP-complete. [If so, PH collapses to 2nd level.]

Today we prove:

theorem: UNSAT $\in$ IP, so coNP $\subseteq$ IP

theorem: #SAT $\in$ IP, so $P^{\#P} \subseteq$ IP

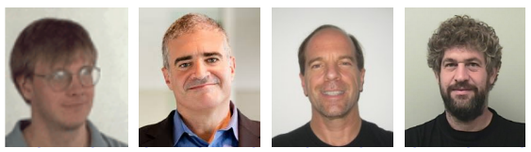
$P^{\#P}$: languages decidable in polynomial time via a machine with a #SAT oracle



These results are surprising:

- Many languages beyond NP.
- The IP for GNI uses properties of graph isomorphisms, but UNSAT/#SAT do not seem to have similar properties.

We learn new ideas:
- ARITHMETIZATION
- SUMCHECK PROTOCOL



Algebraic Methods for Interactive Proof Systems

CARSTEN LUND, LANCE FORTNOW, AND HOWARD KARLOFF

University of Chicago, Chicago, Illinois

AND

NOAM NISAN

Hebrew University, Jerusalem, Israel

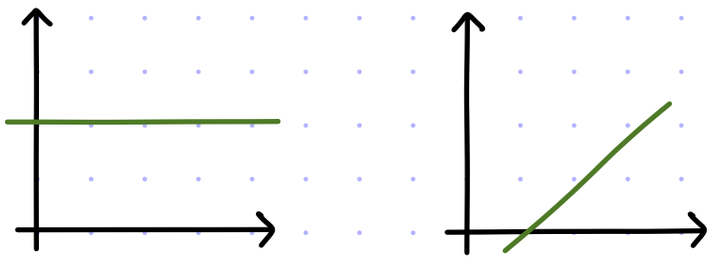
Preliminaries: Zeros of Univariate Polynomials

Basic question: how many zeros can a univariate polynomial p have?

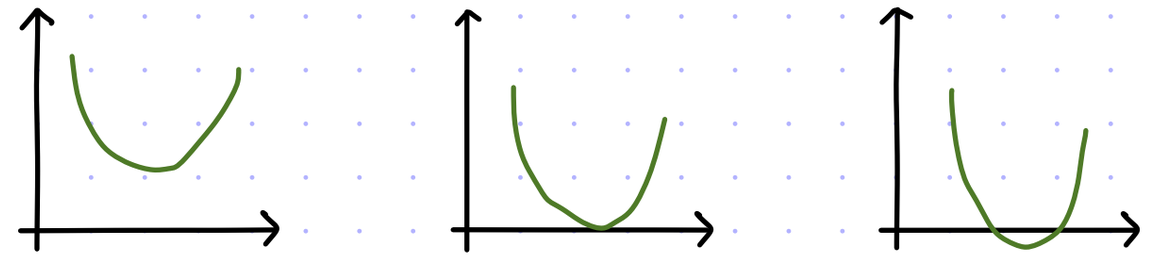
If $p \equiv 0$ then lots.

So assume that $p \not\equiv 0$. In this case the answer depends on the degree of p.

Ex: if p has degree 1
then p has ≤ 1 zeros



Ex: if p has degree 2
then p has ≤ 2 zeros



In general: $p \in \mathbb{F}[x]$ has at most $\deg(p)$ zeros in $\mathbb{F}$
(and exactly $\deg(p)$ zeros in the algebraic closure of $\mathbb{F}$)

This directly leads to an invaluable fact:

Polynomial Identity Lemma: $\forall$ non-zero $f \in \mathbb{F}[x]$ $\forall S \subseteq \mathbb{F}$, $\Pr_{\alpha \leftarrow S}[f(\alpha) = 0] \leq \frac{\deg(f)}{|S|}$

Hence, $\forall f, g \in \mathbb{F}[x]$ if $f \neq g$ then $\Pr_{\alpha \leftarrow S}[f(\alpha) = g(\alpha)] \leq \frac{\max\{\deg(f), \deg(g)\}}{|S|}$.

Preliminaries: Zeros of Multivariate Polynomials

Basic question: how many zeros can a multivariate polynomial p have?

If $p \equiv 0$ then lots.

So assume that $p \not\equiv 0$. In this case the answer depends on the degree of p.

But what do we mean by "degree" of a polynomial $p(x_1,\ldots,x_n) = \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} x_1^{i_1}\cdots x_n^{i_n}$?

- individual degree: $\deg_{ind}(p) := \max\{\max\{i_1,\ldots,i_n\} \mid (i_1,\ldots,i_n) \text{ s.t. } c_{i_1,\ldots,i_n} \neq 0\}$.
- total degree: $\deg_{tot}(p) := \max\{i_1+\cdots+i_n \mid (i_1,\ldots,i_n) \text{ s.t. } c_{i_1,\ldots,i_n} \neq 0\}$.

Examples:
- $p_1 = x_1x_2 + x_1x_4 + x_3$, $\deg_{ind}(p_1) = 1$, $\deg_{tot}(p_1) = 2$
- $p_2 = x_1^3 + x_1x_2 + x_3x_4^2$, $\deg_{ind}(p_2) = 3$, $\deg_{tot}(p_2) = 3$

In general, $\deg_{ind}(p) \le \deg_{tot}(p) \le n \cdot \deg_{ind}(p)$. (Both bounds can be tight.)

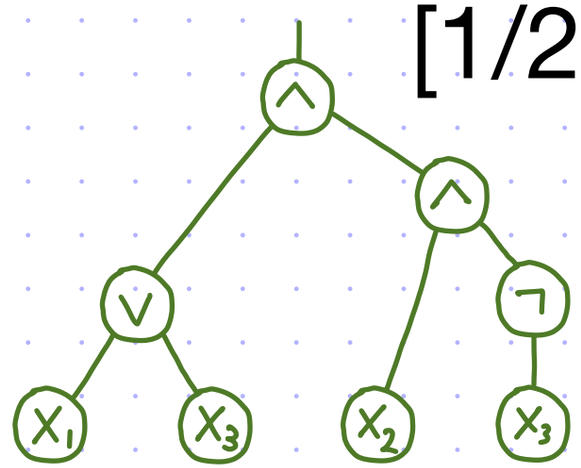
The PIL extends to multivariate polynomials (by induction on n):

$$\forall \text{ non-zero } f \in \mathbb{F}[x_1,\ldots,x_n] \quad \forall S \subseteq \mathbb{F}, \quad \Pr_{\alpha_1,\ldots,\alpha_n \leftarrow S}[f(\alpha_1,\ldots,\alpha_n) = 0] \le \frac{\deg_{tot}(f)}{|S|}.$$

There are refinements and generalizations.

The bound can be tight: $\Pr_{\alpha_1,\ldots,\alpha_n \leftarrow \mathbb{F}}[\prod_{j=1}^{d}(\alpha_1 - \gamma_j) = 0] = \frac{d}{|\mathbb{F}|}$ (for distinct $\gamma_1,\ldots,\gamma_d \in \mathbb{F}$), $\Pr_{\alpha_1,\ldots,\alpha_n \leftarrow \mathbb{F}}[\prod_{i=1}^{n}\alpha_i = 0] = 1 - (1 - \frac{1}{|\mathbb{F}|})^n \ge \frac{n}{|\mathbb{F}|} - \frac{1}{2}\cdot\frac{n^2}{|\mathbb{F}|^2}$, ...

Arithmetization of a Boolean Formula [1/2]



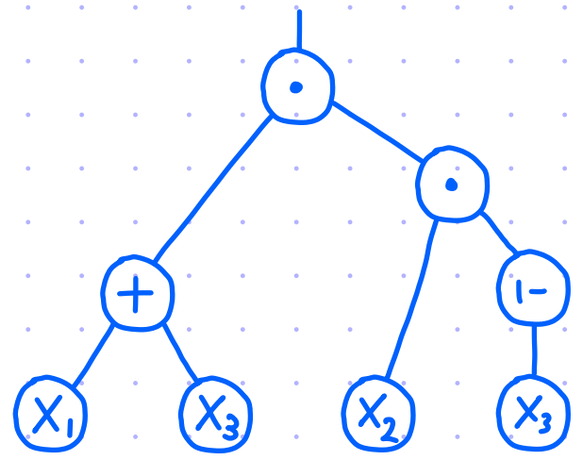
A boolean formula $\varphi(x_1,\dots,x_n)$ is a tree where:

- every leaf vertex is labeled with a variable x_i;
- every internal vertex is a logical operator on its children.

$\neg, \wedge, \vee$

Arithmetization replaces each logical operator with an arithmetic operator:

$$\begin{aligned} \neg x &\mapsto 1-x \\ x \wedge y &\mapsto x \cdot y \\ x \vee y &\mapsto x+y \end{aligned}$$



$$(x_1+x_3)\cdot(x_2\cdot(1-x_3))$$
$$= x_1x_2 - x_1x_2x_3 + x_2x_3 - x_2x_3^2$$

We obtain an expression for a polynomial $p(x_1,\dots,x_n)$ s.t.

- $\deg_{tot}(p) \le$ # leaves in φ
  $\le$ # vertices in φ
  $= |\varphi|$

$\deg_{tot}(x_i) = 1$
$\deg_{tot}(1-p) \le \deg_{tot}(p)$
$\deg_{tot}(p_1+p_2) \le \max\{\deg_{tot}(p_1), \deg_{tot}(p_2)\}$
$\deg_{tot}(p_1\cdot p_2) \le \deg_{tot}(p_1)+\deg_{tot}(p_2)$

- can evaluate the expression for p in $\le |\varphi|$ arithmetic operations:
  # internal vertices in φ $\le$ # vertices in φ $= |\varphi|$

But what does p have to do with φ? Not much in general.

But for 3CNFs we get something useful!

(Also for φ where every NOT is on an input.)

Arithmetization of a Boolean Formula [2/2]

We focus on the case of a **3CNF with m clauses**. 3CNF = conjunctive normal form with 3 literals per clause

It is specified by a subset $S \subseteq [n]^3 \times \{0,1\}^3$ with $|S| = m$:

$$\varphi(x_1,\dots,x_n) = \bigwedge_{\substack{(j_1,j_2,j_3,\\ b_1,b_2,b_3)\in S}} \left(\mathrm{neg}(b_1,x_{j_1}) \vee \mathrm{neg}(b_2,x_{j_2}) \vee \mathrm{neg}(b_3,x_{j_3})\right) .$$

Example: $S = \{(1,3,4,0,1,1),(1,2,5,0,1,0),(3,4,5,1,0,1)\} \mapsto (x_1 \vee \bar{x}_3 \vee \bar{x}_4) \wedge (x_1 \vee \bar{x}_2 \vee x_5) \wedge (\bar{x}_3 \vee x_4 \vee \bar{x}_5)$.

In this case the expression for the polynomial is:

$$p(x_1,\dots,x_n) = \prod_{\substack{(j_1,j_2,j_3,\\ b_1,b_2,b_3)\in S}} \left(\mathrm{neg}(b_1,x_{j_1}) + \mathrm{neg}(b_2,x_{j_2}) + \mathrm{neg}(b_3,x_{j_3})\right) .$$

claim:

- $\varphi \in \mathrm{UNSAT} \rightarrow \sum_{a_1,\dots,a_n\in\{0,1\}} p(a_1,\dots,a_n) = 0$
- $\varphi \notin \mathrm{UNSAT} \rightarrow 0 < \sum_{a_1,\dots,a_n\in\{0,1\}} p(a_1,\dots,a_n) \leq 2^n \cdot 3^m$

corollary: $\forall$ prime $q > 2^n \cdot 3^m$

$$\varphi \in \mathrm{UNSAT} \leftrightarrow \sum_{a_1,\dots,a_n\in\{0,1\}} p(a_1,\dots,a_n) = 0 \mod q$$

Sumcheck Protocol [recursive description]

A protocol for polynomial summations of the form $\sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) = \gamma$.

$\deg_{ind}(p) \le d$

	$P(\mathbb{F}, H, n, \gamma, p)$		$V^p(\mathbb{F}, H, n, \gamma, d)$
Case $n=0$:	Do nothing.		Check that $p = \gamma$.
Case $n>0$:	$p_1(X) := \sum_{\alpha_2,\dots,\alpha_n \in H} p(X, \alpha_2, \dots, \alpha_n)$	$\xrightarrow{p_1 \in \mathbb{F}[X]}$	$\sum_{\alpha_1 \in H} p_1(\alpha_1) \stackrel{?}{=} \gamma$
		$\xleftarrow{\omega_1 \in \mathbb{F}}$	$\omega_1 \leftarrow \mathbb{F}$

Set $p'(X_2,\dots,X_n) := p(\omega_1, X_2, \dots, X_n)$ and $\gamma' := p_1(\omega_1)$.

$P(\mathbb{F}, H, n-1, \gamma', p')$ $\quad \sum_{\alpha_2,\dots,\alpha_n \in H} p'(\alpha_2,\dots,\alpha_n) = \gamma'$ $\quad V^{p'}(\mathbb{F}, H, n-1, \gamma', d)$

Completeness: $\sum_{\alpha_1,\dots,\alpha_n} p(\alpha_1,\dots,\alpha_n) = \gamma \rightarrow \begin{cases} \text{if } n=0 \text{ then } p = \gamma \\ \text{if } n>0 \text{ then } \sum_{\alpha_2,\dots,\alpha_n} p'(\alpha_2,\dots,\alpha_n) = \sum_{\alpha_2,\dots,\alpha_n} p(\omega_1, \alpha_2, \dots, \alpha_n) = p_1(\omega_1) = \gamma' \end{cases}$

Sumcheck Protocol [iterative description]

By "unrolling" the recursion we obtain an iterative description of the protocol.

$P(\mathbb{F}, H, n, \gamma, p)$ $\sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) \stackrel{?}{=} \gamma$ $V^p(\mathbb{F}, H, n, \gamma, d)$ ← $\deg_{ind}(p) \le d$

$p_1(X) := \sum_{\alpha_2,\dots,\alpha_n \in H} p(X, \alpha_2, \dots, \alpha_n)$ — $p_1 \in \mathbb{F}[x]$ → — $\sum_{\alpha_1 \in H} p_1(\alpha_1) \stackrel{?}{=} \gamma$

← $\omega_1 \in \mathbb{F}$ — $\omega_1 \leftarrow \mathbb{F}$

$p_2(X) := \sum_{\alpha_3,\dots,\alpha_n \in H} p(\omega_1, X, \alpha_3, \dots, \alpha_n)$ — $p_2 \in \mathbb{F}[x]$ → — $\sum_{\alpha_2 \in H} p_2(\alpha_2) \stackrel{?}{=} p_1(\omega_1)$

← $\omega_2 \in \mathbb{F}$ — $\omega_2 \leftarrow \mathbb{F}$

⋮ ⋮ ⋮

$p_n(X) := p(\omega_1, \dots, \omega_{n-1}, X)$ — $p_n \in \mathbb{F}[x]$ → — $\sum_{\alpha_n \in H} p_n(\alpha_n) \stackrel{?}{=} p_{n-1}(\omega_{n-1})$

← $\omega_n \in \mathbb{F}$ — $\omega_n \leftarrow \mathbb{F}$

V^p queries p at $(\omega_1, \dots, \omega_n)$ (and makes no other queries)

$p(\omega_1, \dots, \omega_n) \stackrel{?}{=} p_n(\omega_n)$

$O(n \cdot d)$ elements

$O(n \cdot |H| \cdot d)$ field ops + 1 eval of p

Soundness of Sumcheck Protocol [1/2]

claim: $\sum_{\alpha_1,\ldots,\alpha_n \in H} p(\alpha_1,\ldots,\alpha_n) \neq \gamma \longrightarrow \forall \tilde{P}\ \Pr[\langle \tilde{P}, V^p(\mathbb{F},H,n,\gamma,d)\rangle = 1] \leq 1-(1-\frac{d}{|\mathbb{F}|})^n \leq \frac{nd}{|\mathbb{F}|}$.

More generally the bound is $1-\prod_{i\in[n]}(1-\frac{d_i}{|S_i|})$ if $\deg_{x_i}(p)\leq d_i$ and w_i is sampled from $S_i \subseteq \mathbb{F}$.

proof: The malicious prover $\tilde{P}$ is described by n polynomials $\tilde{p}_1,\ldots,\tilde{p}_n \in \mathbb{F}^{\leq d}[X]$ where $\tilde{p}_i$ depends on the verifier messages $w_1,\ldots,w_{i-1} \in \mathbb{F}$.

The proof is by induction on n.

- **Base case: $n=1$.** We show that $\Pr[\langle \tilde{P}, V^p(\mathbb{F},H,n=1,\gamma,d)\rangle = 1] \leq 1-(1-\frac{d}{|\mathbb{F}|}) = \frac{d}{|\mathbb{F}|}$.

The malicious prover $\tilde{P}$ sends a single message $\tilde{p}_1 \in \mathbb{F}^{\leq d}[X]$.

Assume that $\sum_{\alpha_1\in H} \tilde{p}_1(\alpha_1) = \gamma$, since if not the verifier immediately rejects.

Hence $\tilde{p}_1 \neq p_1$ because $\sum_{\alpha_1\in H} p_1(\alpha_1) = \sum_{\alpha_1\in H} p(\alpha_1) \neq \gamma$.

$P(\mathbb{F},H,n=1,\gamma,p)$: $p_1(X) := p(X)$

$\sum_{\alpha_1\in H} p(\alpha_1) \stackrel{?}{=} \gamma$

$p_1 \in \mathbb{F}[X]$ →

← $w_1 \in \mathbb{F}$

$V^p(\mathbb{F},H,n=1,\gamma,d)$: $\sum_{\alpha_1\in H} p_1(\alpha_1) \stackrel{?}{=} \gamma$; $w_1 \leftarrow \mathbb{F}$; $p(w_1) \stackrel{?}{=} p_1(w_1)$

We conclude that

$$\Pr[\langle \tilde{P}, V^p(\mathbb{F},H,n=1,\gamma,d)\rangle = 1] = \Pr[p(w_1) = \tilde{p}_1(w_1)] = \Pr[p_1(w_1) = \tilde{p}_1(w_1)] \leq \frac{d}{|\mathbb{F}|}.$$

Soundness of Sumcheck Protocol [2/2]

claim: $\sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) \neq \gamma \longrightarrow \forall\, \tilde{P}\ \Pr[\langle \tilde{P}, V^p(\mathbb{F},H,n,\gamma,d)\rangle = 1] \leq 1-(1-\frac{d}{|\mathbb{F}|})^n \leq \frac{nd}{|\mathbb{F}|}$.

proof: [continued]

- Inductive case: $n>1$. Assume the claim for $(n-1)$-variate polynomials.

We can assume that $\sum_{\alpha_1 \in H} \tilde{p}_1(\alpha_1) = \gamma$ since otherwise the verifier immediately rejects.

Hence $\tilde{p}_1 \neq p_1$ because $\sum_{\alpha_1 \in H} p_1(\alpha_1) = \sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) \neq \gamma$.

Define $E_1 := \text{“}\tilde{p}_1(w_1) \neq \sum_{\alpha_2,\dots,\alpha_n \in H} p(w_1,\alpha_2,\dots,\alpha_n)\text{”}$.

Then $\Pr[E_1] = \Pr_{w_1 \leftarrow \mathbb{F}}[\tilde{p}_1(w_1) \neq p_1(w_1)] \geq 1-\frac{d}{|\mathbb{F}|}$.

Define $E_2 := \text{“the verifier rejects in round } i \geq 2\text{”}$.

By the induction hypothesis, $\Pr[E_2 | E_1] \geq (1-\frac{d}{|\mathbb{F}|})^{n-1}$.

We conclude that

$$\begin{aligned}
&\Pr[\langle \tilde{P}, V^p(\mathbb{F},H,n,\gamma,d)\rangle = 1] \\
&= 1 - \Pr[E_2] \\
&\leq 1 - \Pr[E_2|E_1]\cdot\Pr[E_1] \\
&\leq 1 - \left(1-\frac{d}{|\mathbb{F}|}\right)^{n-1}\cdot\left(1-\frac{d}{|\mathbb{F}|}\right) = 1-\left(1-\frac{d}{|\mathbb{F}|}\right)^n.
\end{aligned}$$

■

Interactive Proof for UNSAT

We describe an IP for $\mathrm{UNSAT} = \{\varphi \mid \varphi$ is an unsatisfiable 3CNF boolean formula$\}$.

This implies that $\mathrm{coNP} \subseteq \mathrm{IP}$ (by reducing to UNSAT via polynomial-time reductions).

$P(\varphi)$

$p := \mathrm{ARITH}(\varphi, \mathbb{F}_q)$

$P_{sc}(\mathbb{F}_q, \{0,1\}, n, 0, p)$
field domain #vars sum polynomial

$q \in \mathbb{N}$

sumcheck protocol for $\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} p(\alpha_1,\dots,\alpha_n) = 0$

$V(\varphi)$

$2^n \cdot 3^m < q < 2^{\mathrm{poly}(m,n)}$

$q \in \mathrm{PRIMES}$

$p := \mathrm{ARITH}(\varphi, \mathbb{F}_q)$

$V_{sc}^{p}(\mathbb{F}_q, \{0,1\}, n, 0, \deg_{ind}(p))$
field domain #vars sum ind degree

$(w_1,\dots,w_n) \in \mathbb{F}_q^n \rightsquigarrow p(w_1,\dots,w_n) \in \mathbb{F}_q$
poly(m,n) time

Completeness: $\varphi \in \mathrm{UNSAT} \to \sum_{\alpha_1,\dots,\alpha_n} p(\alpha_1,\dots,\alpha_n) = 0 \to$ P always convinces V.

Soundness: $\varphi \notin \mathrm{UNSAT} \to \sum_{\alpha_1,\dots,\alpha_n} p(\alpha_1,\dots,\alpha_n) \neq 0 \to$ error is $\leq \frac{n \cdot \deg_{ind}(p)}{q} \leq \frac{n \cdot \deg_{tot}(p)}{q} \leq \frac{n \cdot m}{q} < \frac{n \cdot m}{2^n 3^m}$.

Arithmetization for #SAT

The arithmetization we used for UNSAT was coarse:

$$\forall (a_1,\dots,a_n)\in\{0,1\}^n \qquad \varphi(a_1,\dots,a_n)=0 \;\to\; p(a_1,\dots,a_n)=0$$
$$\varphi(a_1,\dots,a_n)=1 \;\to\; 0<p(a_1,\dots,a_n)\le 3^m$$

We can modify the arithmetization to be more precise:

$$\begin{aligned} \neg x &\mapsto 1-x \\ x\wedge y &\mapsto x\cdot y \\ x\vee y &\mapsto x+y-x\cdot y \end{aligned}$$

[Explanation: $x\vee y=\overline{\bar{x}\wedge\bar{y}} \to 1-(1-x)\cdot(1-y)=x+y-x\cdot y$]

For every boolean formula φ: $\deg_{tot}(p)\le|\varphi|$, can evaluate p in $\le O(|\varphi|)$ operations, and p agrees with φ on every boolean input (even if φ not a 3CNF).

claim: $\forall (a_1,\dots,a_n)\in\{0,1\}^n \quad \varphi(a_1,\dots,a_n)=0 \to p(a_1,\dots,a_n)=0$, $\varphi(a_1,\dots,a_n)=1 \to p(a_1,\dots,a_n)=1$

p is a low-degree extension of φ because $p|_{\{0,1\}^n}\equiv\varphi$

We can now reduce #SAT to a sumcheck problem:

corollary: $\forall$ prime $q>2^n$ $\quad \#\varphi=c \iff \sum_{a_1,\dots,a_n\in\{0,1\}} p(a_1,\dots,a_n)=c \bmod q$

Interactive Proof for #SAT

We describe an IP for $\#SAT = \{(\varphi, c) \mid \varphi \text{ is a boolean formula with } c \text{ satisfiable assignments}\}$.

$P((\varphi, c))$

$V((\varphi, c))$

$\xrightarrow{q \in \mathbb{N}}$

$2^n < q < 2^{poly(m,n)}$

$q \in PRIMES$

$p := \text{ARITH}^{\#}(\varphi, \mathbb{F}_q)$

$p := \text{ARITH}^{\#}(\varphi, \mathbb{F}_q)$

$P_{sc}(\mathbb{F}_q, \{0,1\}, n, c, p)$
field domain #vars sum polynomial

sumcheck protocol for $\sum_{\alpha_1,\dots,\alpha_n \in \{0,1\}} p(\alpha_1,\dots,\alpha_n) = c$

$V_{sc}^{p}(\mathbb{F}_q, \{0,1\}, n, c, deg_{ind}(p))$
field domain #vars sum ind degree

$(w_1,\dots,w_n) \in \mathbb{F}_q^n \rightsquigarrow p(w_1,\dots,w_n) \in \mathbb{F}_q$

poly(m,n) time

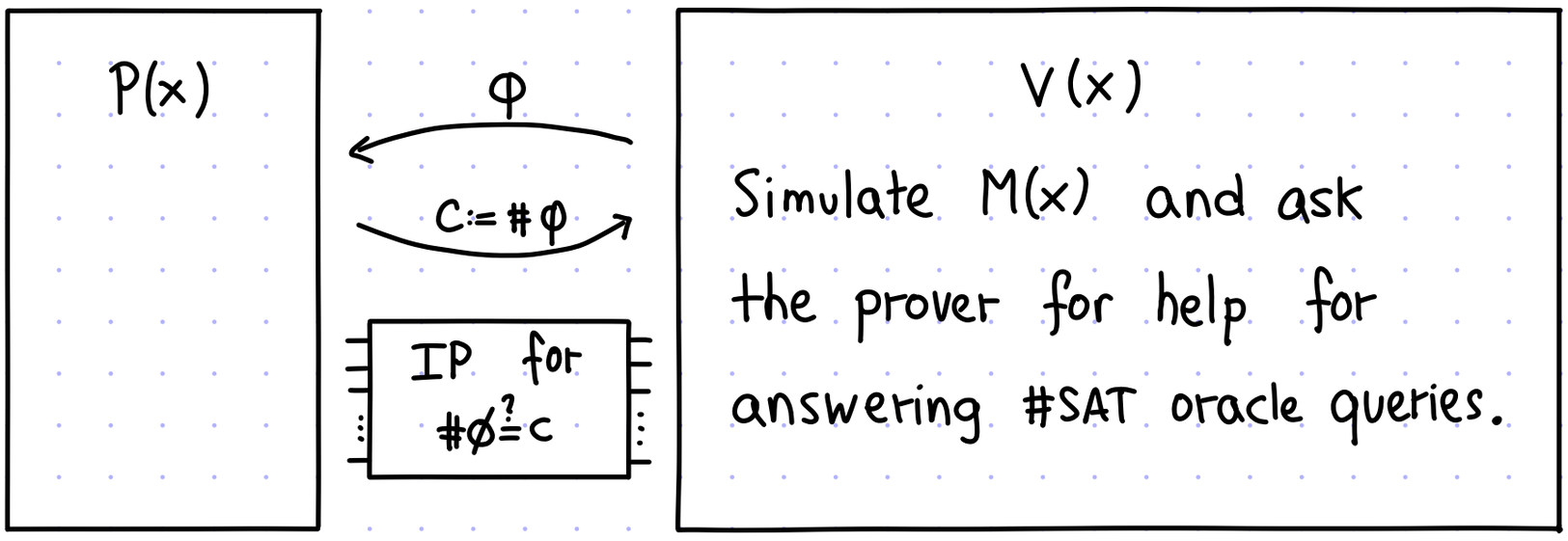
Completeness: $(\varphi, c) \in \#SAT \to \sum_{\alpha_1,\dots,\alpha_n} p(\alpha_1,\dots,\alpha_n) = c \to$ P always convinces V.

Soundness: $(\varphi, c) \notin \#SAT \to \sum_{\alpha_1,\dots,\alpha_n} p(\alpha_1,\dots,\alpha_n) \neq c \to$ error is $\leq \frac{n \cdot deg_{ind}(p)}{q} \leq \frac{n \cdot deg_{tot}(p)}{q} \leq \frac{n \cdot 3m}{q} < \frac{n \cdot 3m}{2^n}$.

Interactive Proof for $P^{\#P}$

theorem: $P^{\#P} \subseteq IP$

proof: Let $L \in P^{\#P}$, and let M be a machine that decides L with a #SAT oracle. We describe an IP for L.



Completeness: $x \in L \rightarrow \Pr[\langle P(x), V(x)\rangle = 1] = 1$ because: $M^{\#SAT}(x) = 1$; $P(x)$ answers correctly each #SAT call; and the IP for #SAT has perfect completeness.

Soundness: $x \notin L \rightarrow \forall \tilde{P}\ \Pr[\langle \tilde{P}, V(x)\rangle = 1] \leq \varepsilon_0$ where ε_0 = "soundess error of the IP for #SAT". Indeed, $M^{\#SAT}(x) = 0$ so $M(x) = 1$ in the simulation happens only if $\tilde{P}$ lies on a #SAT query, in which case we rely on the soundness of the IP for #SAT. ■

History of Arithmetization

Arithmetization loosely refers to useful ways to map problems in boolean logic to problems that involve arithmetic.

$\{0,1\}, \Sigma \rightarrow \mathbb{N}, \mathbb{Q}, \mathbb{F}$

logical ops $\rightarrow$ arithmetic ops

- [Gödel 1931]: encode a string $(a_1,\dots,a_n) \in \Sigma^n$ as a number $p_1^{a_1} \cdots p_n^{a_n} \in \mathbb{N}$ ($\text{PRIMES} = \{p_1, p_2, \dots\}$)
  This Gödel numbering enables establishing the incompleteness of formal arithmetic.
- [Church 1936]: uses ↯ to prove that a specific problem in number theory is undecidable
  (Subsequently, Turing showed how to do this via Turing machines.)

→ Convenient injection from strings to natural numbers.

- [Razborov 1987] [Smolensky 1987]: map boolean circuit to a low-degree polynomial approximation
  This enables proving circuit complexity lower bounds. (E.g. $\text{PARITY} \notin AC_0$.)

← poly(n)-size O(1)-depth circuits with any fan-in

→ Polynomials of low degree cannot describe certain functions.

- [Lund Fortnow Karloff Nisan 1992]: $\text{PERMANENT} \in IP$ by viewing the computation of the permanent as a sumcheck claim
- [Shamir 1992]: more boolean logic → polynomial arithmetic
- ... huge role in the probabilistic proof literature

efficient encoding/decoding
multiplication property
rich automorphism group
local testing/decoding/correction
...

→ Polynomials of low degree are codes with MANY useful properties.

Bonus: Sumcheck Protocol as a Reduction

The sumcheck protocol can be phrased as a **reduction** (from a polynomial summation to a single evaluation):

$$\begin{cases} \sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) = \gamma \longrightarrow \Pr\left[p(w_1,\dots,w_n)=v \text{ where } ((w_1,\dots,w_n),v) \leftarrow \langle P(\mathbb{F},H,n,\gamma,p), V(\mathbb{F},H,n,\gamma,d)\rangle \right] = 1 \\ \sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) \neq \gamma \longrightarrow \forall \tilde{P}\ \Pr\left[p(w_1,\dots,w_n)=v \text{ where } ((w_1,\dots,w_n),v) \leftarrow \langle \tilde{P}, V(\mathbb{F},H,n,\gamma,d)\rangle \right] \leq 1-\left(1-\frac{d}{|\mathbb{F}|}\right)^n \end{cases}$$

$P(\mathbb{F},H,n,\gamma,p)$ $\sum_{\alpha_1,\dots,\alpha_n \in H} p(\alpha_1,\dots,\alpha_n) \stackrel{?}{=} \gamma$ $V^{\blacksquare}(\mathbb{F},H,n,\gamma,d)$ } NO query access to p

$p_1(X) := \sum_{\alpha_2,\dots,\alpha_n \in H} p(X,\alpha_2,\dots,\alpha_n)$	$\xrightarrow{p_1 \in \mathbb{F}[x]}$	$\sum_{\alpha_1 \in H} p_1(\alpha_1) \stackrel{?}{=} \gamma$
	$\xleftarrow{w_1 \in \mathbb{F}}$	$w_1 \leftarrow \mathbb{F}$
$p_2(X) := \sum_{\alpha_3,\dots,\alpha_n \in H} p(w_1,X,\alpha_3,\dots,\alpha_n)$	$\xrightarrow{p_2 \in \mathbb{F}[x]}$	$\sum_{\alpha_2 \in H} p_2(\alpha_2) \stackrel{?}{=} p_1(w_1)$
	$\xleftarrow{w_2 \in \mathbb{F}}$	$w_2 \leftarrow \mathbb{F}$
$\vdots$	$\vdots$	$\vdots$
$p_n(X) := p(w_1,\dots,w_{n-1},X)$	$\xrightarrow{p_n \in \mathbb{F}[x]}$	$\sum_{\alpha_n \in H} p_n(\alpha_n) \stackrel{?}{=} p_{n-1}(w_{n-1})$
	$\xleftarrow{w_n \in \mathbb{F}}$	$w_n \leftarrow \mathbb{F}$
		Output $((w_1,\dots,w_n), p_n(w_n))$

We use this perspective throughout the course.

Bibliography

Sumcheck protocol

- [LFKN 1992]: Algebraic methods for interactive proof systems, by Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan.
- The unreasonable power of the sumcheck protocol, by Justin Thaler.

Polynomial identity lemma

- [Schwartz 1980]: Fast probabilistic algorithms for verification of polynomial identities, by Jacob Schwartz.
- [Zippel 1979]: Probabilistic algorithms for sparse polynomials, by Richard Zippel.
- [BCPS 2015]: On zeros of a polynomial in a finite grid, by Anurag Bishnoi, Pete Clark, Aditya Potukuchi, John R. Schmitt.
- The curious history of the Schwartz–Zippel lemma, by Richard Lipton.

Arithmetization

- [Razborov 1987]: Lower bounds on the size of bounded depth circuits over a complete basis with logical addition, by Alexander Razborov.
- [Smolensky 1987]: Algebraic methods in the theory of lower bounds for boolean circuit complexity, by Roman Smolensky.

Efficient sumcheck prover

- [BDT 2024]: The sumcheck protocol over fields of small characteristic, by Suyash Bagad, Yuval Domb, Justin Thaler.
- [BDDT 2025]: Speeding up sumcheck proving, by Suyash Bagad, Quang Dao, Yuval Domb, Justin Thaler.
- [BCFFMMZ 2025]: Time-space tradeoffs for sumcheck, by Anubhav Baweja, Alessandro Chiesa, Elisabetta Fedele, Giacomo Fenzi, Pratyush Mishra, Tushar Mopuri, Andrew Zitek-Estrada.