

# The proof theoretic strength of the Steinitz Exchange Theorem

Michael Soltys  
McMaster University, Canada

Full version of this paper in *Discrete Applied Mathematics*.

A **propositional proof system (PPS)** is just a **PolyTime** function  $f : \Sigma^* \rightarrow \Sigma^*$ , where  $\text{range}(f) = \text{TAUTOLOGIES}$ .

$x \in \Sigma^*$  is a **proof** of  $\tau$  iff  $f(x) = \tau$ .

A PPS  $f$  is **poly bounded** iff there exists a polynomial  $p$  such that for all  $\tau$ ,  $\exists x \in \Sigma^*$ , such that  $|x| \leq p(|\tau|)$  and  $f(x) = \tau$ .

A poly bounded PPS exists  $\iff \mathbf{NP} = \mathbf{co-NP}$ .

An **automatizable**, poly bounded, PPS exists  $\iff \mathbf{P} = \mathbf{NP}$ .

*Context: complexity theory, automated reasoning, “reverse mathematics”, ...*

- **Open Question:** is there a poly bounded PPS?
- **Program:** show that stronger and stronger PPS are not poly bounded.
- **Next Step** in the Program: show a separation between Frege and Extended Frege.

**Frege:** lines are boolean *formulas*.

**Extended Frege:** lines are boolean *circuits*.

(Allow abbreviations in proofs:  $a \equiv \phi$ )

Candidates for the separation:

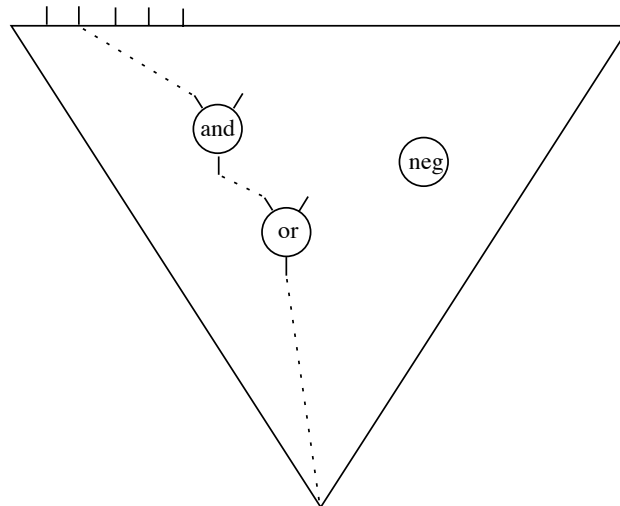
*universal principles of matrix algebra,*

for example:

- $AB = I \rightarrow BA = I$
- $p_A(A) = 0$ ,  $p_A$  is the *char poly* of  $A$ , Cayley-Hamilton Thm
- $\det(AB) = \det(A) \det(B)$

Start with the following aim:

Show “**NC<sup>2</sup> concepts**” prove the Cayley-Hamilton Thm



We encode a set of vectors  $\{v_1, v_2, \dots, v_n\}$  as a matrix

$$T = [v_1 v_2 \dots v_n].$$

**Steinitz Exchange Theorem (SET):** if  $T$  is *total*, and  $E$  is *linearly independent*, then there exists  $F \subseteq T$ , such that  $|F| = |E|$ , and  $(T - F) \cup E$  is total.

$$\exists X, TX = I \wedge (\forall Y \neq 0, EY \neq 0) \rightarrow \exists F \subseteq T, |F| = |E| \wedge \exists X, (T - F \cup E)X = I$$

So **SET** is a  $\Pi_2^B$  formulas of **QLA**.

Given  $T, E$ , the matrix  $F$  can be computed in  $\mathbf{NC}^2$ .

Suppose  $E = [e_1 e_2]$  and  $T = [t_1 t_2 t_3 t_4]$ , then consider

$$[e_1 e_2] \quad [e_1 e_2 t_1] \quad [e_1 e_2 t_1 t_2] \quad [e_1 e_2 t_1 t_2 t_3] \quad [e_1 e_2 t_1 t_2 t_3 t_4]$$

*Independently* for every  $i = 0, 1, 2, 3$ , if

$$\text{Rank}([e_1 e_2 t_1 \dots t_i]) = \text{Rank}([e_1 e_2 t_1 \dots t_{i+1}])$$

then put  $t_{i+1}$  in  $F$ .

Rank can be computed in  $\mathbf{NC}^2$  with *Mulmuley's algorithm*.

## Mulmuley's Algorithm<sup>a</sup>

$M$  is  $n \times n$  matrix,  $p_M(x)$  its char. poly. (Berkowitz's alg.)

*Geometric Rank* : usual rank.

*Algebraic Rank* :  $n - (\text{highest power of } x \text{ that divides } p_M(x))$ .

---

<sup>a</sup>*Parallel Linear Algebra*, by Joachim von zur Gathen, chapter in *Synthesis of Parallel Algorithms*.

## Mulmuley's Algorithm<sup>a</sup>

$M$  is  $n \times n$  matrix,  $p_M(x)$  its char. poly. (Berkowitz's alg.)

*Geometric Rank* : usual rank.

*Algebraic Rank* :  $n - (\text{highest power of } x \text{ that divides } p_M(x))$ .

**Claim:**  $\text{Rank}_G(M) = \text{Rank}_G(M^2) \Rightarrow \text{Rank}_G(M) = \text{Rank}_A(M)$ .

---

<sup>a</sup>*Parallel Linear Algebra*, by Joachim von zur Gathen, chapter in *Synthesis of Parallel Algorithms*.

## Mulmuley's Algorithm<sup>a</sup>

$M$  is  $n \times n$  matrix,  $p_M(x)$  its char. poly. (Berkowitz's alg.)

*Geometric Rank* : usual rank.

*Algebraic Rank* :  $n - (\text{highest power of } x \text{ that divides } p_M(x))$ .

**Claim:**  $\text{Rank}_G(M) = \text{Rank}_G(M^2) \Rightarrow \text{Rank}_G(M) = \text{Rank}_A(M)$ .

Given  $M$ , let  $M'$  be

$$\begin{pmatrix} 0 & M^t \\ M & 0 \end{pmatrix}$$

Clearly,  $\text{Rank}_G(M) = \frac{1}{2} \text{Rank}_G(M')$ .

---

<sup>a</sup>*Parallel Linear Algebra*, by Joachim von zur Gathen, chapter in *Synthesis of Parallel Algorithms*.

## Mulmuley's Algorithm<sup>a</sup>

$M$  is  $n \times n$  matrix,  $p_M(x)$  its char. poly. (Berkowitz's alg.)

*Geometric Rank* : usual rank.

*Algebraic Rank* :  $n - (\text{highest power of } x \text{ that divides } p_M(x))$ .

**Claim:**  $\text{Rank}_G(M) = \text{Rank}_G(M^2) \Rightarrow \text{Rank}_G(M) = \text{Rank}_A(M)$ .

Given  $M$ , let  $M'$  be

$$\begin{pmatrix} 0 & M^t \\ M & 0 \end{pmatrix}$$

Clearly,  $\text{Rank}_G(M) = \frac{1}{2} \text{Rank}_G(M')$ .

$M'' = M' \cdot \text{diag}(1, y, y^2, \dots, y^{2n-1})$ ;  $\text{Rank}_G(M'') = \text{Rank}_G((M'')^2)$ .

$y$  is an indeterminate, and  $M''$  is a matrix over the field  $\mathbb{F}(y)$ .

---

<sup>a</sup>*Parallel Linear Algebra*, by Joachim von zur Gathen, chapter in *Synthesis of Parallel Algorithms*.

**SET** can be shown with **PolyTime** concepts.

Can it be shown with **NC<sup>2</sup>** concepts?

**SET** proves in **QLA** the following principles

1.  $(\exists B \neq 0)[AB = I \vee AB = 0]$ ,
2. The columns of an  $n \times (n + 1)$  matrix are linearly dependent,
3. Every matrix has an annihilating polynomial,
4.  $AB = I \supset BA = I$ ,
5. **Existence of  $A^n$ , and the Cayley-Hamilton Thm.**

**QLA** can prove the existence of powers of a matrix from **SET**.

Let  $\text{POW}(A, n)$  be the formula:

$$\exists \langle X_0 X_1 \dots X_n \rangle (\forall i \leq n) [X_0 = I \wedge (i < n \supset X_{i+1} = X_i * A)]$$

Show **QLA**  $\vdash (\exists B \neq 0) [AB = I \vee AB = 0] \supset \text{POW}(A, n)$ .

Let  $N$  be the  $n^2 \times n^2$  matrix consisting of  $n \times n$  blocks which are all zero except for  $(n - 1)$  copies of  $A$  above the diagonal zero blocks<sup>a</sup>.

---

<sup>a</sup>*A taxonomy of problems with fast parallel algorithms*, by Stephen Cook.

Let  $N$  be the  $n^2 \times n^2$  matrix consisting of  $n \times n$  blocks which are all zero except for  $(n - 1)$  copies of  $A$  above the diagonal zero blocks<sup>a</sup>. Then  $N^n = 0$ , and  $(I - N)^{-1} = I + N + N^2 + \dots + N^{n-1} =$

$$\begin{pmatrix} I & A & A^2 & \dots & A^{n-1} \\ 0 & I & A & \dots & A^{n-2} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I \end{pmatrix}.$$

Set  $C = I - N$ .

---

<sup>a</sup> *A taxonomy of problems with fast parallel algorithms*, by Stephen Cook.

Let  $N$  be the  $n^2 \times n^2$  matrix consisting of  $n \times n$  blocks which are all zero except for  $(n - 1)$  copies of  $A$  above the diagonal zero blocks<sup>a</sup>. Then  $N^n = 0$ , and  $(I - N)^{-1} = I + N + N^2 + \dots + N^{n-1} =$

$$\begin{pmatrix} I & A & A^2 & \dots & A^{n-1} \\ 0 & I & A & \dots & A^{n-2} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I \end{pmatrix}.$$

Set  $C = I - N$ .

Show that if  $CB = 0$ , then  $B = 0$ , using induction on the rows of  $B$ , starting with the bottom row.

---

<sup>a</sup> *A taxonomy of problems with fast parallel algorithms*, by Stephen Cook.

Let  $N$  be the  $n^2 \times n^2$  matrix consisting of  $n \times n$  blocks which are all zero except for  $(n - 1)$  copies of  $A$  above the diagonal zero blocks<sup>a</sup>. Then  $N^n = 0$ , and  $(I - N)^{-1} = I + N + N^2 + \dots + N^{n-1} =$

$$\begin{pmatrix} I & A & A^2 & \dots & A^{n-1} \\ 0 & I & A & \dots & A^{n-2} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I \end{pmatrix}.$$

Set  $C = I - N$ .

Show that if  $CB = 0$ , then  $B = 0$ , using induction on the rows of  $B$ , starting with the bottom row.

Using  $(\exists B \neq 0)[CB = I \vee CB = 0]$ , conclude that there is a  $B$  such that  $CB = I$ . Finally, show that  $B = I + N + N^2 + \dots + N^{n-1}$ .

<sup>a</sup>*A taxonomy of problems with fast parallel algorithms*, by Stephen Cook.

## Strong Linear Independence (SLI)

if  $\{v_1, \dots, v_m\}$  are  $n \times 1$ , non-zero, linearly dependent vectors, then there exists a  $1 \leq k < m$  such that

$$\begin{array}{c} \text{lin. indep.} \\ \underbrace{\{v_1, \dots, v_k, v_{k+1}, v_{k+2}, \dots, v_m\}} \\ \text{lin. dep.} \end{array}$$

Csanky's algorithm ( $\mathbf{NC}^2$ ) for computing the characteristic poly. of a matrix uses Newton's symmetric polynomials:

$$s_0 = 1,$$

$$s_k = \frac{1}{k} \sum_{i=1}^k (-1)^{i-1} s_{k-i} \operatorname{tr}(A^i)$$

$$p_A(x) := s_0 x^n - s_1 x^{n-1} + s_2 x^{n-2} - \dots \pm s_n x^0.$$

**Theorem:** QLA proves the **Cayley-Hamilton Thm.** from SET and SLI.

*The 12 steps proof :*

- (1)  $p_A$  is the characteristic polynomial of the matrix  $A$  as computed by Csanky's algorithm.
- (2) Let  $W = \{e_i, Ae_i, \dots, A^n e_i\}$ .
- (3) By SET,  $W$  must be linearly dependent.
- (4) By SLI there exists a  $k \leq n$  such that  $W_0 = \{e_i, Ae_i, \dots, A^{k-1} e_i\}$  is linearly independent and  $k$  is the largest such index.

(5)  $A^k e_i$  can be written as a linear combination of the vectors in  $W_0$ .

Let  $c_1, \dots, c_k$  be the coefficients of this linear combination, so that if  $g(x) = x^k + c_1 x^{k-1} + \dots + c_k$ , then  $g(A)e_i = 0$ .

(6) Let  $A_g$  be the  $k \times k$  *companion matrix* of  $g$ ,

$$\left( \begin{array}{c|cccc} 0 & 0 & 0 & \dots & 0 & -c_k \\ \hline 1 & 0 & 0 & \dots & 0 & -c_{k-1} \\ 0 & 1 & 0 & \dots & 0 & -c_{k-2} \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & -c_1 \end{array} \right)$$

(7) **LAP** proves  $p_{A_g} = g$ , and so **LAP** proves  $(p_{A_g}(A))e_i = 0$ .

(8) Extend  $W_0$  to  $B = W_0 \cup \{e_{j_1}, \dots, e_{j_{n-k}}\}$ .

Existence of  $B$  follows from **SET**:

let  $T = B_0 =$  the standard basis,

let  $E = W_0$ , which is linearly independent,

let  $F = B_0 - \{e_{j_1}, \dots, e_{j_{n-k}}\}$ ,

so  $B = (T - F) \cup E$ .

$$A \sim \begin{pmatrix} A_g & E_1 \\ 0 & E_2 \end{pmatrix}$$

(9) **LAP** proves that if  $C_1 \sim C_2$  then  $p_{C_1}(x) = p_{C_2}(x)$   
( $\text{tr}(A) = \text{tr}(PAP^{-1})$ , since  $\text{tr}(AB) = \text{tr}(BA)$ ).

(10) **LAP** proves that if

$$C = \begin{pmatrix} C_1 & * \\ 0 & C_2 \end{pmatrix}$$

then  $p_C(x) = p_{C_1}(x) \cdot p_{C_2}(x)$ .

(11)  $\therefore p_A(A)e_i = (p_{A_g}(A) \cdot p_E(A))e_i = p_E(A) \cdot (p_{A_g}(A)e_i) = 0$ .

(12) This is true for all  $e_i$  in the standard basis, and so  $p_A(A) = 0$ .

**Question:**

Can algorithms be proven correct *within* the complexity classes that they run in ?