

# The Proof Complexity of Linear Algebra

Michael Soltys and Stephen Cook

July 2002

What is the complexity of proving the fundamental principles of linear algebra?

For example:

- $AB = I \supset BA = I$
- The Cayley-Hamilton Theorem,  
 $p_A(A) = 0$
- The multiplicativity of determinant,  
 $\det(AB) = \det(A) \det(B)$

$$AB = I \supset BA = I$$

Proposed by Cook as a candidate for separating Frege and Extended Frege.

It turns out to be  $p$ -equivalent in Frege (over a fixed field such as  $\mathbb{Z}_p$  or  $\mathbb{Q}$ ) to:

- $(AB = I \wedge AC = I) \supset B = C$
- $AB = I \supset (AC \neq 0 \vee C = 0)$
- $AB = I \supset BA = I$
- $AB = I \supset A^t B^t = I$

$AB = I \supset BA = I$  can be proven in polysize Extended Frege, because Gaussian Elimination can be proven correct in polysize Extended Frege.

Correct means: *any matrix  $A$  can be converted to row-echelon form with a sequence of elementary row operations.*

Using this condition of correctness, we can show that if  $AB = I$ , then there exists a left inverse  $C$  of  $A$ . Since it can be proven in polysize Frege that  $AB = I \supset A(BA - I) = 0$ , we are done.

On the other hand, it seems that  $AB = I \supset BA = I$  cannot be proven with polysize Frege proofs.

Gaussian Elimination is a well studied polytime algorithm, whose correctness can be shown easily in polysize Extended Frege.

However, it seems to be inherently sequential.

There are fast parallel algorithms for computing the characteristic polynomial (and hence the inverse) of matrices:

- Berkowitz's algorithm
- Chistov's algorithm
- Csanky's algorithm (for fields of char 0)

All three algorithms can be formalized with  $NC^2$  circuits (circuits of polynomial size, and depth  $O(\log^2)$ , in the size of matrices).

We concentrate on Berkowitz's algorithm, because:

- It does not use divisions, so it is field independent (so it works over commutative rings).
- It is easy to formalize, since it only requires matrix powering as a primitive operation. This is an advantage because we want to design logical theories for proving matrix identities.

**Question:** Can we prove correctness conditions for Berkowitz's algorithm using  $NC^2$  concepts only, just as we were able to prove correctness conditions for Gaussian Elimination using polytime concepts only?

What would be the desirable correctness conditions for Berkowitz's algorithm?

Berkowitz's algorithm computes the coefficients of the characteristic polynomial of a matrix, using iterated matrix products.

Two fundamental properties related to the characteristic polynomial, are:

- The Cayley-Hamilton Theorem,  
 $p_A(A) = 0$ ,
- multiplicativity of determinant,  
 $\det(AB) = \det(A) \det(B)$   
(the constant coefficient of the char poly of an  $n \times n$  matrix  $A$  is  $(-1)^n \det(A)$ ).

In fact,  $AB = I \supset BA = I$  follows (in polysize  $\text{NC}^2$ -Frege) from either property, and hence all the other matrix identities, presumably hard for Frege, also follow (in polysize  $\text{NC}^2$ -Frege) from them.

**So the question is:**

Can we prove the Cayley-Hamilton Theorem and the multiplicativity of the determinant in polysize  $\text{NC}^2$ -Frege?

To study this problem, and to see what type of reasoning is necessary to formalize the universal theorems of linear algebra, we developed three logical theories:

$$\text{LA} \subseteq \text{LAP} \subseteq \forall \text{LAP}$$

of increasing strength.

## Linear Algebra (LA)

LA is a quantifier-free theory, with three sorts: indices, field elements, matrices.

It contains:

- the usual axioms of arithmetic for indices ( $\mathbb{N}$ ), and induction on indices
- the usual field axioms
- a way of constructing new matrices (using some rudimentary  $\lambda$ -calculus).

For example, given  $n \times n$  matrices  $A, B$ , we construct the new matrix  $A + B$  as  $\lambda ij \langle n, n, A_{ij} + B_{ij} \rangle$ .

All the usual ring properties of matrices can be proven in LA. For example:

$$A(BC) = (AB)C$$

$$A + B = B + A$$

However, this is a fairly weak theory; all theorems can be translated to families of tautologies, with uniform polysize Frege proofs.

The underlying field ( $\mathbb{Z}_p$  or  $\mathbb{Q}$ ) is a parameter in the translation, even though the theories LA, LAP, and  $\forall$ LAP, are field independent.

Also, LA proves the equivalence of the matrix identities that we proposed as candidates for separating Frege and Extended Frege.

**OPEN PROBLEM:** Is  $AB = I \supset BA = I$  (and hence the other identities) independent of LA?

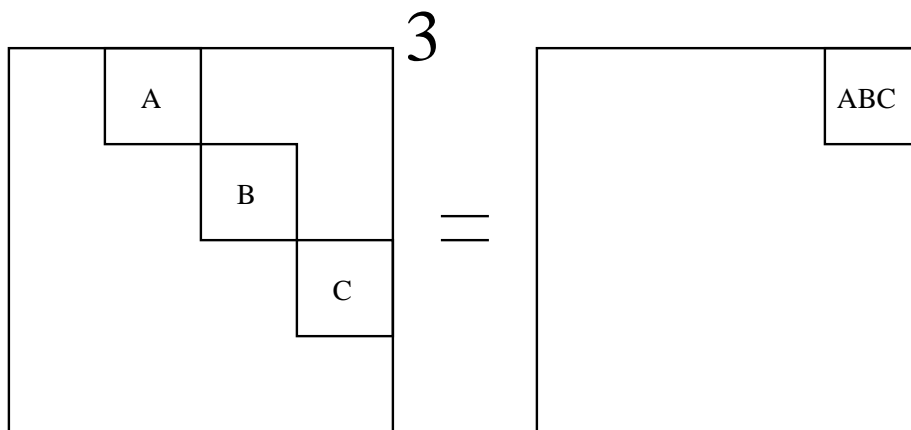
## Linear Algebra with matrix powers (LAP)

We extend the theory LA to LAP by adding a new function symbol,  $P$ , with the following defining axioms:

$$P(A, 0) = I$$

$$P(A, n + 1) = P(A, n) * A$$

We can compute iterated matrix product with powering, for example, we can compute  $ABC$  as follows:



Since we can compute the  $n$ -th power of a matrix by repeated squaring in  $\log n$  many steps, it follows that we can compute iterated matrix products with  $NC^2$  circuits.

In LAP we can express Berkowitz's algorithm. Berkowitz's algorithm computes the char polynomial of  $A$  in terms of the char polynomial of  $M$ .

$$A = \begin{array}{|c|c|} \hline a_{11} & R \\ \hline S & M \\ \hline \end{array}$$

What connects the char poly of  $A$  to the char poly of  $M$  is **Samuelson's identity**:

$$p_A(x) = (x - a_{11})p_M(x) - R * \text{adj}(xI - M) * S$$

Since  $\text{adj}(xI - M)$  can be expressed in terms of the coefficients of  $p_M$ , and using the Cayley-Hamilton Theorem, we can restate Samuelson's identity as a matrix identity.

For example, if  $A$  is a  $3 \times 3$  matrix, then:

$$p_A = \begin{pmatrix} 1 & 0 & 0 \\ -a_{11} & 1 & 0 \\ -RS & -a_{11} & 1 \\ -RMS & -RS & -a_{11} \end{pmatrix} p_M$$

where  $p_A$  and  $p_M$  are column vectors containing the coefficients of the char polys of  $A$  and  $M$ , respectively.

Thus  $p_A = C_1 p_M$ , where  $C_1$  is a Toeplitz lower triangular matrix, and  $p_A, p_M$  are the char polys of  $A, M$ , respectively.

We can repeat this procedure to compute  $p_M$ , by defining  $p_M$  in terms of the char poly of  $M[1|1]$ . Continuing this way we get:

$$p_A = C_1 C_2 \cdots C_n$$

where each  $C_j$  is define in terms of  $M_j, R_j, S_j$ :

$$A = \begin{array}{|c|} \hline \dots \\ \hline \begin{array}{|c|c|} \hline a_{jj} & R_j \\ \hline S_j & M_j \\ \hline \end{array} \\ \hline \end{array}$$

Thus, we can compute the coefficients of the characteristic polynomial of a matrix in LAP, with Berkowitz's algorithm. Furthermore:

**THEOREM:** LAP proves the equivalence of the following principles:

- The Cayley-Hamilton Theorem.
- The axiomatic definition of determinant (alternating, multilinear, and 1 on  $I$ ).
- The cofactor expansion formula.

and LAP also proves that the Cayley-Hamilton Theorem follows from the multiplicativity of the determinant.

Furthermore,  $AB = I \supset BA = I$  (and the other matrix identities) follows in LAP from *any* of the above principles.

**OPEN PROBLEM:** Can any of these principles be proven in LAP?

We outline the proof of the axiomatic definition of the determinant from the Cayley-Hamilton Theorem.

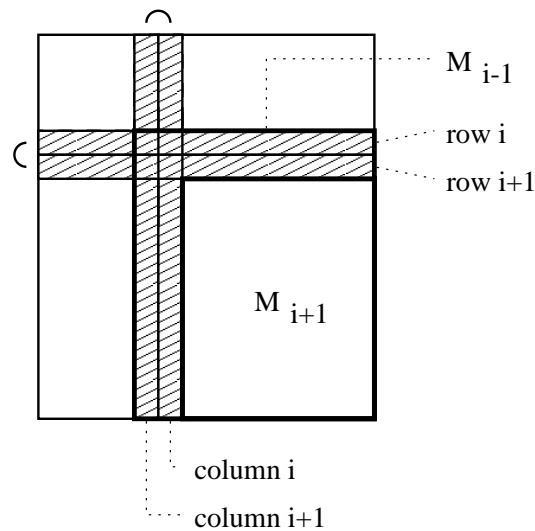
–It is easy to show that  $\det(I_n) = 1$  (induction on  $n$ , noting that the principal submatrix of  $I_n$  is  $I_{n-1}$ ).

–It is easy to show multilinearity in the first row, directly from Berkowitz's algorithm. To show multilinearity in all rows, we need alternation.

–Alternation is the difficult part. The proof of alternation has the following components:

(i) Let  $M_i$  be the  $i$ -th principal submatrix of  $A$  (so  $M_1 = A[1|1]$ , and  $M_{i+1} = M_i[1|1]$ ). Show (in LAP) that if the C-H Theorem holds for  $M_i$ , then:

$A$  and  $I_{i(i+1)} A I_{i(i+1)}$  have the same char poly.



(ii) Using the C-H Theorem several times, show that:

$A$  and  $I_{ij}AI_{ij}$

have the same char poly.

(iii) Show, using the C-H Theorem on  $M_2$ , show that:

$$\det(A) = -\det(I_{12}A)$$

(iv) Combine (ii) and (iii) to show that:

$$\det(A) = -\det(I_{ij}A)$$

The theorems of LA can be translated to feasible polysize Frege proofs.

The theorems of LAP can be translated to feasible quasi-polysize Frege proofs, i.e., into  $NC^2$ -Frege proofs.

Thus, the equivalence of the principles mentioned in the previous slide can be shown with polysize  $NC^2$ -Frege proofs.

## $\forall$ LAP

This is LAP where we allow induction on formulas of the type  $(\forall X \leq n)\alpha$ , where  $\alpha$  has no quantifiers, and  $X$  is a variable of type matrix of size at most  $n \times n$ .

This strengthening of induction is what allows us to prove the Cayley-Hamilton Theorem.

The theorems of  $\forall$ LAP can be translated to uniform polysize Extended Frege proofs.

Therefore, our proof of the Cayley-Hamilton Theorem is a feasible proof; in fact, as far as we know, this is the first feasible proof of this principle.

All previous proofs of the Cayley-Hamilton Theorem relied on the Lagrange Expansion of the determinant, which for a matrix of size  $n$  has  $n!$  terms, and hence it is very infeasible.

The idea behind the  $\forall$ LAP proof of the C-H Theorem is the following:

If  $p_A(A) \neq 0$ , that is, if the C-H theorem fails for  $A$ , then we can find *in polytime* a sub-matrix  $B$  of  $A$  for which  $p_B(B) \neq 0$ , i.e., for which the C-H theorem fails already.

Since the C-H Theorem does *not* fail for  $1 \times 1$  matrices, after at most  $n = (\text{size of } A)$  steps we get a contradiction.

This idea can be expressed with universal quantifiers over variables of type matrix: if the C-H theorem holds for all matrices smaller than  $A$ , then it also holds for  $A$ .

## Summary of Propositional Translations

Theory	Propositional Proof System (and corresponding complexity class)	
LA	field: $\mathbb{Z}_p$	field: $\mathbb{Q}$
	polysize BD Frege with MOD $p$ gates ( $AC^0[p]$ )	polysize Frege ( $NC^1$ )
LAP	quasi-polysize Frege ( $DET \subseteq NC^2$ )	
$\forall$ LAP	polysize Extended Frege (P)	

## Open Problems

- Is  $AB = I \supset BA = I$  independent of LA?
- Can the Cayley-Hamilton Theorem and/or the multiplicativity of the determinant be shown in LAP?
- Can  $AB = I \supset BA = I$  be shown in LAP?
- More generally, can universal matrix identities be shown with quasi-polysize Frege proofs?