The Proof Complexity of Linear Algebra

Michael Soltys
Department of Computing and Software
McMaster University
1280 Main Street West
Hamilton, Ontario L8S4K1, CANADA
Email: <soltys@mcmaster.ca>

Stephen Cook
Department of Computer Science
University of Toronto
10 King's College Road
Toronto, Ontario M5S 3G4, CANADA
Email: <sacook@cs.toronto.edu>

Abstract

We introduce three formal theories of increasing strength for linear algebra in order to study the complexity of the concepts needed to prove the basic theorems of the subject. We give what is apparently the first feasible proofs of the Cayley-Hamilton theorem and other properties of the determinant, and study the propositional proof complexity of matrix identities.

1 Introduction

The complexity of the basic operations of linear algebra such as the determinant and matrix inverse have been well-studied. Over the field of rationals they lie within the complexity class NC^2, and are complete for the class DET [6]. Here we are concerned with the *proof complexity* of linear algebra, which roughly speaking is the complexity of the concepts needed to prove the basic properties of these operations. In general proof complexity has two aspects: uniform and nonuniform (see [7] for a treatise on the subject). The uniform aspect concerns the power of logical theories required to prove a given assertion, while the nonuniform aspect concerns the power of propositional proof systems required to yield polynomial size proofs of a tautology family representing the assertion.

The method of Gaussian elimination can be used to give polynomial time algorithms for the determinant, matrix inverse, etc., but it does not yield the fast parallel algorithms which place these operations in NC^2. We base our treatment of linear algebra on Berkowitz's elegant algorithm [2], which gives field-independent reductions of these operations to matrix powering (the complexity class DET) (see [10] for alternative algorithms).

We are interested in the question of whether the basic properties of the determinant can be proved using concepts restricted to the class DET, and we make this question precise by defining a quantifier-free theory LAP formalizing reasoning about matrix algebra based on matrix powering. We use LAP to present Berkowitz's algorithm. Since this algorithm computes not only the determinant of a given square matrix A, but also the coefficients of the characteristic polynomial $p_A(x) = \det(xI - A)$, it is natural to ask whether LAP proves the Cayley-Hamilton (C-H) theorem, which asserts $p_A(A) = 0$. We leave this question open, but we demonstrate its importance by showing that LAP proves the equivalence of the C-H theorem with two other basic results: the cofactor expansion of the determinant and the axiomatic definition of the determinant.

If we cannot prove the C-H theorem in LAP, can we at least find a feasible proof; i.e., one using only polynomial time concepts? This question (over finite fields and over the rationals) has a natural precise formalization, since feasible reasoning has been well-studied using $\forall$-equivalent theories such as Cook's PV [5] or Buss's S^1_2 [4]. A study of the linear algebra literature has turned up no such feasible proof, and in fact most proofs of the C-H theorem are based directly or indirectly on the Lagrange expansion of the determinant, which represents an exponential time algorithm.

Thus a major contribution of this paper is our success in finding a feasible proof of the C-H theorem. We formalize this proof in the field-independent theory $\forall$LAP, which extends LAP by allowing induction over formulas with bounded universal matrix quantifiers. We justify the label "feasible" for the proof in several ways, including an interpretation of $\forall$LAP (when the underlying field is finite or the rationals) into the feasible theory V^1_1 (equivalent to Buss's S^1_2). Our feasible proof yields feasible proofs of many basic matrix properties, including the multiplicativity of the determinant, and the correctness of algorithms based on Gaussian elimination.

One specific motivation for this research is to find natural tautology families which may distinguish the power of Frege and Extended Frege (eFrege) propositional proof systems. (A line in a Frege proof is a propositional formula which is an immediate logical consequence of earlier lines, whereas a line in an eFrege proof may also introduce a new propositional variable by definition, allowing for concise abbreviations of exponentially long formulas). The principle

$$AB = I \Longrightarrow BA = I \tag{1}$$

where A and B are $n \times n$ matrices, may provide such an example. This principle (over $\mathbb{Z}_2$ or $\mathbb{Z}$) is readily translated into a tautology INV_n of size polynomial in n. We conjecture that the family $\langle \text{INV}_n \rangle$ does not have polynomial size Frege proofs, since the proof of (1) seems to require concepts such as Gaussian elimination or matrix powering whose complexity apparently cannot be expressed by polynomial size propositional formulas (i.e., is not in NC^1). On the other hand, we show that (1) can be proved using polynomial time concepts, and hence $\langle \text{INV}_n \rangle$ does have polynomial size eFrege proofs.

Altogether we introduce three logical theories of increasing power

$$\mathsf{LA} \subset \mathsf{LAP} \subset \forall\mathsf{LAP}$$

to formalize linear algebra reasoning. Each theory has three sorts: indices (i.e., natural numbers), field elements, and matrices, and all theorems hold for any choice of the underlying field. The base theory LA allows the basic ring properties of matrices to be formulated and proved. The principle (1) can be formulated in LA but (we conjecture) not proved. We show that LA proves the equivalence of (1) with other "hard" matrix identities. Theorems of LA translate into tautology families with polynomial size Frege proofs.

We extend LA to LAP by adding a new function, P, which is intended to denote matrix powering, i.e., $\mathsf{P}(n, A)$ means A^n. LAP is well suited for formalizing Berkowitz's algorithm, and it is strong enough to prove the equivalence of some fundamental principles of linear algebra. The theorems of LAP translate into quasi-poly-bounded Frege proofs.

We finally extend LAP to $\forall$LAP by allowing induction on formulas with bounded universal matrix quantifiers. This new theory is strong enough to prove the C-H theorem, and hence (by our equivalence) all the major principles of Linear Algebra. The theorems of $\forall$LAP translate into poly-bounded Extended Frege proofs.

This paper is based on the PhD thesis [8] of the first author, which is available on the Web.

2 The Theory LA

We define a quantifier-free theory of Linear Algebra (matrix algebra), and call it LA. Our theory is strong enough to prove the ring properties of matrices such as $A(BC) = (AB)C$ and $A + B = B + A$ but weak enough so that all the theorems of LA (over finite fields or the field of rationals) translate into propositional tautologies with short Frege proofs.

Our theory has three sorts of object: *indices* (i.e., natural numbers), *field elements*, and *matrices*, where the corresponding variables are denoted $i, j, k, ...$; $a, b, c, ...$; and $A, B, C, ...$, respectively. The semantics assumes that objects of type field are from a fixed but arbitrary field, and objects of type matrix have entries from that field.

Terms and formulas are built from the function and predicate symbols:

$$0_{\text{index}}, 1_{\text{index}}, +_{\text{index}}, *_{\text{index}}, -_{\text{index}}, \mathtt{div}, \mathtt{rem}, 0_{\text{field}}, 1_{\text{field}}, +_{\text{field}}, *_{\text{field}}, -_{\text{field}}, {}^{-1}\, \mathtt{r}, \mathtt{c}, \mathtt{e}, \Sigma, \leq_{\text{index}}, =_{\text{index}}, =_{\text{field}}, =_{\text{matrix}}, \text{cond}_{\text{index}}, \text{cond}_{\text{field}} \tag{2}$$

The intended meanings should be clear, except for the following operations on a matrix A: $\mathtt{r}(A), \mathtt{c}(A)$ are the numbers of rows and columns in A, $\mathtt{e}(A, i, j)$ is the field element A_{ij}, $\Sigma(A)$ is the sum of the elements in A. Also $\text{cond}(\alpha, t_1, t_2)$ is interpreted **if** α **then** t_1 **else** t_2, where α is a formula all of whose atomic sub-formulas have the form $m \leq n$ or $m = n$, where m, n are terms of type index, and t_1, t_2 are terms either both of type index or both of type field. The subscripts $_{\text{index}}$ and $_{\text{field}}$ are usually omitted, since they are clear from the context.

Atomic formulas and formulas are built in the usual manner, except no quantifiers are allowed.

We use Gentzen's sequent calculus LK (with quantifier rules omitted) for the underlying logic. We include 34 non-logical axioms in four groups: Axioms for equality, indices, field elements, and matrices (all quantifier-free). These specify the basic properties of the function and predicate symbols (2). By convention each instance of an axiom resulting from substituting terms for variables is also an axiom, so the axioms are really axiom schemes.

We need just two non-logical rules: an equality rule for terms of type matrix, and the induction rule:

$$\frac{\Gamma, \alpha(i) \rightarrow \alpha(i+1), \Delta}{\Gamma, \alpha(0) \rightarrow \alpha(n), \Delta} \tag{3}$$

In addition to the usual rules for constructing terms we also allow the terms $\lambda ij\langle m, n, t\rangle$ of type matrix. Here i and j are variables of type index bound by the λ operator, intended to range over the rows and columns of the matrix. Here also m, n are terms of type index *not* containing i, j (representing the numbers of rows and columns of the matrix) and t is a term of type field (representing the matrix element in position (i, j)).

The λ terms allow us to construct the sum, product, transpose, etc., of matrices. For example, suppose first that

A and B are $m \times n$ matrices. Then, $A + B$ can be defined as $\lambda ij\langle m, n, \mathsf{e}(A, i, j) + \mathsf{e}(B, i, j)\rangle$. Now suppose that A and B are $m \times p$ and $p \times n$ matrices, respectively. Then:

$$A * B := \lambda ij\langle m, n, \Sigma\lambda kl\langle p, 1, \mathsf{e}(A, i, k) * \mathsf{e}(B, k, j)\rangle\rangle$$

However, even if matrices are of incompatible size, their addition and product is well defined, since the "smaller" matrix is implicitly padded with zeros (as $\mathsf{e}(A, i, j) = 0$ for i or j outside the range). Thus, all terms are well defined.

3 The Theorems of LA

We show that all matrix identities which state that the set of $n \times n$ matrices form a ring, and all identities that state that the set of $m \times n$ matrices form a module over the underlying field, are theorems of LA. However, LA is apparently not strong enough to prove matrix identities which require arguing about inverses. Here are four examples (stated as Gentzen sequents), which we refer to as *hard* matrix identities:

$$AB = I, AC = I \to B = C \qquad \text{I}$$
$$AB = I \to AC \neq 0, C = 0 \qquad \text{II}$$
$$AB = I \to BA = I \qquad \text{III}$$
$$AB = I \to A^tB^t = I \qquad \text{IV}$$

(where A^t is the transpose of A.). We show that LA proves them all equivalent, but conjecture that none is provable in LA. In section 6 we show that these are theorems of the stronger theory $\forall\mathsf{LAP}$. We speculate that $\forall\mathsf{LAP}$ might be conservative over LA + α, where α is any of I–IV.

In sections 5 we show that the theorems of LA translate into tautology families with polynomial size Frege proofs, and later we argue that theorems of $\forall\mathsf{LAP}$ translate into tautology families with polynomial size eFrege proofs. We conjecture that the translations of I–IV do not have polynomial size Frege proofs. This conjecture is partly inspired by the paper [3] in which the "Odd Town Theorem" was presented as a candidate combinatorial principle for separating Frege and eFrege systems, since its proof seems to require an independence argument from linear algebra. We show that the "Odd Town Theorem" can be formulated in LA and follows in LA from any of the principles I–IV, and hence its propositional translations have polynomial size eFrege proofs. We are unable to show in LA that conversely these hard matrix identities follow from the "Odd Town Theorem."

Throughout this paper we prove properties of matrices by induction on their size. We outline briefly this technique. The size of a matrix has two parameters: the number of rows, and the number of columns. Suppose that we want to prove that some property holds for all matrices A. We define a new (constructed) matrix $M(i, A)$, and we let $d(A)$ be defined as:

$$d(A) := \text{cond}(r(A) \leq \mathsf{c}(A), \mathsf{r}(A), \mathsf{c}(A))$$

that is, $d(A) = \min\{\mathsf{r}(A), \mathsf{c}(A)\}$. Now let:

$$\begin{aligned} M(i, A) := \lambda pq\langle &\mathsf{r}(A) - d(A) + i, \mathsf{c}(A) - d(A) + i, \\ &\mathsf{e}(A, d(A) - i + p, d(A) - i + q)\rangle \end{aligned}$$

that is, $M(i, A)$ is the i-th principal sub-matrix of A. For example, if A is a 3×5 matrix, then $M(1, A)$ is a 1×3 matrix, with the entries from the lower-right corner of A.

To prove that a property $\mathcal{P}$ holds for A, we prove that $\mathcal{P}$ holds for $M(1, A)$ (the Basis Case), and then we prove that if $\mathcal{P}$ holds for $M(i, A)$, then $\mathcal{P}$ also holds for $M(i + 1, A)$ (the Induction Step). From this we conclude, by the induction rule, that $\mathcal{P}$ holds for $M(d(A), A)$, and $M(d(A), A)$ is just A. Note that in the Basis Case we might have to prove that $\mathcal{P}$ holds for a row vector or a column vector, which is a $k \times 1$ or a $1 \times k$ matrix, and this in turn can also be done by induction (on k).

4 Berkowitz's Algorithm and LAP

Berkowitz's algorithm allows us to reduce the computation of the characteristic polynomial $p_A(x) = \det(xI - A)$ of an $n \times n$ matrix A to the operation of matrix powering. Suppose

$$A = \begin{pmatrix} a_{11} & R \\ S & M \end{pmatrix} \qquad (4)$$

where R is an $1 \times (n - 1)$ row matrix and S is a $(n - 1) \times 1$ column matrix and M is $(n-1) \times (n-1)$. Let $p(x)$ and $q(x)$ be the characteristic polynomials of A and M respectively. Suppose that the coefficients of p form the column vector

$$p = \begin{pmatrix} p_n & p_{n-1} & \dots & p_0 \end{pmatrix}^t \qquad (5)$$

where p_i is the coefficient of x^i in $\det(xI - A)$, and similarly for q. Then Berkowitz showed

$$p = C_1 q \qquad (6)$$

where C_1 is an $(n + 1) \times n$ Toeplitz lower triangular matrix (Toeplitz means that the values on each diagonal are the same) and where the entries in the first column are defined as follows:

$$c_{i1} = \begin{cases} 1 & \text{if } i = 1 \\ -a_{11} & \text{if } i = 2 \\ -(RM^{i-3}S) & \text{if } i \geq 3 \end{cases} \qquad (7)$$

Berkowitz's algorithm consists in repeating this for q, and continuing so that p is expressed as a product of matrices:

$$p = C_1 C_2 \cdots C_n \qquad (8)$$

where C_i is an $(n+2-i) \times (n+1-i)$ Toeplitz matrix defined as in (7) except A is replaced by its i-th principal sub-matrix.

Since each element of C_i can be explicitly defined in terms of A using matrix powering, and since the iterated matrix product can be reduced to matrix powering by a standard method, the entire product (8) can be expressed in terms of A using matrix powering.

To formalize Berkowitz's algorithm we extend the theory LA to the theory LAP by adding a new function symbol P, where $\mathsf{P}(n, A)$ means A^n. We also add two new axioms, which give a recursive definition of P; namely, $\mathsf{P}(0, A) = I$ and $\mathsf{P}(n+1, A) = \mathsf{P}(n, A) * A$.

Thus the right-hand side of (8) can be expressed as a term in LAP. We use this term as the definition in LAP of the characteristic polynomial p, given in (5), of the matrix A. (If $n = 1$ and $A = (a)$, then $p = (1 \quad -a)^t$.)

Also in LAP we define

$$\det(A) := (-1)^n p_0 \tag{9}$$

where p_0 is as in (5), and we define

$$\text{adj}(A) := (-1)^{n-1}(p_n A^{n-1} + p_{n-1} A^{n-2} + \ldots + p_1 I) \tag{10}$$

Recall that in the usual definition, the (i, j)-th entry of the adjoint of A is $(-1)^{i+j}\det(A[i|j])$, where $A[i|j]$ is the minor obtained by deleting the i-th row and j-th column of A. The equivalence of this and (10) can be proved in LAP using the Cayley-Hamilton (C-H) Theorem as an assumption.

Recall that the C-H theorem states that $p(A) = 0$. From (10) we have that:

$$A \text{ adj}(A) = (-1)^{n-1}(p(A) - p_0 I)$$

Assuming $p(A) = 0$ we have by (9) that:

$$A \text{ adj}(A) = \text{adj}(A) A = \det(A) I \tag{11}$$

In fact LAP easily proves the equivalence of (11) with the C-H theorem. We also have

Theorem 4.1 LAP proves that the C-H theorem implies the hard matrix identities I–IV of section 3.

Proof.(Outline) It suffices to consider the identity III:

$$AB = I \rightarrow BA = I$$

From the assumption $AB = I$ it suffices to show that there is *some* left inverse C of A, since using simple ring properties of matrices (formalizable in LA) it is easy to show $AB = I$ and $CA = I$ implies $BA = I$.

To show that a left inverse C exists, we use the C-H theorem $p(A) = 0$, where p is the characteristic polynomial of A. Since p is not the zero polynomial (it has leading coefficient 1), there must be $k \geq 0$ and a polynomial q such that

$$0 = p(A) = q(A) A^k \tag{12}$$

where q has a nonzero constant term. From $AB = I$ we can show in LAP by induction on i that $A^i B^i = I$. Thus multiplying (12) on the right by B^k we obtain $q(A) = 0$, which we can write as

$$\hat{q}(A) A = -q_0 I$$

where q_0 is the constant coefficient of q. Dividing by $-q_0$ we obtain the required left inverse $C = (-1/q_0)\hat{q}(A)$. □

It is an open question whether LAP proves the C-H theorem in general, although we show that it proves the C-H theorem for triangular matrices.

By the *axiomatic definition of the determinant* we mean that the determinant function $\det(A)$ satisfies the three conditions

- $\det$ is multi-linear in the rows and columns of A
- $\det$ is alternating in the rows and columns of A
- if $A = I$, then $\det(A) = 1$

It is well-known that these conditions completely characterize the determinant.

By the *cofactor expansion* we mean

$$1 \leq i \leq n, \ \det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A[i|j]) \tag{13}$$

where $A[i|j]$ denotes the matrix obtained from A by removing the i-th row and the j-th column. For each i, the RHS of the equation is called the *cofactor expansion of A along the i-th row*, and (13) states that we obtain $\det(A)$ expanding along any row of A. Applying this recursively results in an exponential time algorithm for computing $\det(A)$, showing that the expansion completely defines the determinant.

By the *multiplicativity of the determinant* we mean

$$\det(AB) = \det(A)\det(B)$$

where A, B are $n \times n$ matrices.

The following is the major result of this section.

Theorem 4.2 LAP proves the equivalence of each of the following:

1. C-H theorem
2. Axiomatic definition of det
3. Cofactor Expansion

and LAP also proves the following implications:

1. Multiplicativity of det $\Longrightarrow$ C-H theorem
2. C-H Theorem + $\{\det(A) = 0 \rightarrow AB \neq I\}$
   $\Longrightarrow$ Multiplicativity of det.

The rest of section 4 will consist of an outline of the proof of this theorem, given in sections 4.1, 4.2, 4.3, and 4.4 (all the details can be found in [8, Chapter 6]). Later, in section 6, we will give feasible proofs (i.e. proofs involving only polynomial time concepts) of the C-H theorem and of $\det(A) = 0 \rightarrow AB \neq I$. From this it will follow that all the results mentioned in Theorem 4.2 have feasible proofs, and from other results mentioned in section 6 it will follow that the propositional translations of these results have polynomial size Extended Frege proofs.

The following result is used frequently in the proof: LAP proves

$$\text{det}(A) = a_{11}\text{det}(M) - R\text{adj}(M)S \tag{14}$$

where A is given by (4). The proof is straightforward from the definitions involved.

4.1 The axiomatic definition of determinant

We show that when the determinant is defined as in (9), the axiomatic definition of the determinant follows from the C-H theorem, and that this can be proven in LAP. The condition $\det(I) = 1$ is easy, and multilinearity in the first row (and column) is easy as well. Thus the whole proof hinges on an LAP proof of alternation from the C-H theorem.

It is in fact enough to prove alternation in the rows, as alternation in the columns will follow from alternation in the rows by $\det(A) = \det(A^t)$—which can be derived in LAP by induction on the size of A (see [8, Lemma 5.1.7]).

In order to show alternation, we define I_{ij} to be the matrix obtained from the identity matrix by interchanging the i-th and j-th rows. The effect of multiplying A on the left by I_{ij} is that of interchanging the i-th and j-th rows of A. On the other hand, AI_{ij} is A with the i-th and j-th columns interchanged. We sometimes abbreviate $I_{i,i+1}$ by I_i.

We show alternation in the rows by first showing that for any matrix A, A and I_1AI_1 have the same char poly ($I_1 = I_{1,2}$, so I_1AI_1 is the matrix A with the first two rows interchanged, and the first two columns interchanged). Then, we show that A and I_iAI_i have the same char poly for any i ($I_i = I_{i,i+1}$). Finally, we obtain that A and $I_{ij}AI_{ij}$ have the same char poly (as any permutation is a product of transpositions).

We also show that $\det(A) = -\det(I_1A)$. From this it follows that $\det(A) = -\det(I_{1i}A)$ for all i, since we can bring the i-th row to the second position (via $I_{2i}AI_{2i}$), and reorder things (by applying $I_{2i}AI_{2i}$ once more). Since $I_{ij} = I_{1i}I_{1j}I_{1i}$, this gives us alternation in the rows.

Note that we prove that A and $I_{ij}AI_{ij}$ have the same char poly, i.e., $p_{I_{ij}AI_{ij}} = p_A$, to be able to reorder the matrix and prove alternation.

4.2 The cofactor expansion

We show that LAP proves that the cofactor expansion formula (13) follows from the axiomatic definition of the determinant. We first show that the cofactor expansion of A along the first row is equal to $\det(A)$. Define A_j, for $1 \leq j \leq n$, to be A, with the first row replaced by zeros, except for the $(1, j)$-th entry which remains unchanged. Then, using multilinearity along the first row of A, we obtain:

$$\det(A) = \det(A_1) + \det(A_2) + \cdots + \det(A_n) \tag{15}$$

Consider A_j, for $j > 1$. If we interchange the first column and the j-th column, and then, with $(j - 2)$ transpositions we bring the first column (which is now in the j-th position) to the second position, we obtain, by alternation and (14), the following:

$$\begin{aligned}\det(A_j) &= (-1)^{j-1}a_{1j}\det(A[1|j]) \\ &= (-1)^{1+j}a_{1j}\det(A[1|j])\end{aligned}$$

From this, and from equation (15), we obtain the cofactor expansion along the first row, that is, we obtain (13) for $i = 1$.

If we want to carry out the cofactor expansion along the i-th row (where $i > 1$), we interchange the first and the i-th row, and then we bring the first row (which is now in the i-th position) to the second row with $(i - 2)$ transposition. Denote this new matrix A', and note that $\det(A') = (-1)^{i-1}\det(A)$. Now, expanding along the first row of A', we obtain (13) for $i > 1$.

4.3 The adjoint as a matrix of cofactors

We wish to show that LAP proves the C-H theorem from the cofactor expansion formula (i.e., from (13)). To this end, we first show that (13) implies (in LAP) the axiomatic definition of determinant.

We want to show that we can get multilinearity, alternation and $\det(I) = 1$ from (13). To show multilinearity along row (column) i, we just expand along row (column) i. To show $\det(I) = 1$ use induction on the size of I; in fact, showing that $\det(I) = 1$ can be done in LAP without any assumptions.

It is very easy to show that alternation follows from multilinearity and from:

If two rows (columns) of A are equal $\Longrightarrow \det(A) = 0$

To show this in LAP (from the cofactor expansion formula), we expand along row i first to obtain:

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(A[i|k])$$

and then we expand each minor $A[i|k]$ along the row that corresponds to the j-th row of A. Note that we end up with $n(n-1)$ terms; polynomially many in the size of A. Since row i is identical to the row j, we can pair each term with its negation; hence the result is zero, so $\det(A) = 0$.

Therefore, we have that the axiomatic definition of the determinant follows from the cofactor expansion formula, in LAP. We can now proceed, and finish showing the equivalences in Theorem 4.2, by showing that the cofactor expansion formula implies the C-H theorem, also in LAP.

We start by showing that LAP proves that:

$$\mathrm{adj}(A) = ((-1)^{i+j} \det(A[j|i]))_{ij}$$

i.e., that $\mathrm{adj}(A)$ is the transpose of the matrix of cofactors of A, from the axiomatic definition of det.

Consider the following matrix:

$$C = \begin{pmatrix} 0 & e_i^t \\ e_j & A \end{pmatrix}$$

where e_i is a column vector with zeros everywhere except in the i-th position where it has a 1. By (14), we have that:

$$\det(C) = -e_i^t \mathrm{adj}(A) e_j = \text{ (i,j)-th entry of } -\mathrm{adj}(A)$$

On the other hand, from alternation on C, we have that $\det(C) = (-1)^{i+j+1} \det(A[j|i])$. To see this, note that we need $(j+1)$ transpositions to bring the j-th row of A to the first row in the matrix C, to obtain the following matrix:

$$C' = \begin{pmatrix} 1 & A_j \\ 0 & e_i^t \\ 0 & A[j|-] \end{pmatrix}$$

where A_j denotes the j-th row of A, and $A[j|-]$ denotes A with the j-th row deleted. Then, by (14), we have:

$$\det(C') = \det \begin{pmatrix} e_i^t \\ A[j|-] \end{pmatrix}$$

and now with i transpositions, we bring the i-th column of $\begin{pmatrix} e_i^t \\ A[j|-] \end{pmatrix}$ to the first column, to obtain: $\begin{pmatrix} 1 & 0 \\ 0 & A[j|i] \end{pmatrix}$. Therefore, $\det(C') = (-1)^i \det(A[j|i])$ finishing the proof.

Therefore, LAP proves that the (i,j)-th entry of $\mathrm{adj}(A)$ is given by $(-1)^{i+j} \det(A[j|i])$.

Note that $p_A(A) = 0$ can also be stated as $A\mathrm{adj}(A) = \det(A)I$, using our definitions of the adjoint and the determinant. Thus, the following shows that LAP proves the C-H theorem from the cofactor expansion formula: LAP proves $A\mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)I$ from the cofactor expansion formula.

We show first that $A\mathrm{adj}(A) = \det(A)I$. The (i,j)-th entry of $A\mathrm{adj}(A)$ is equal to:

$$a_{i1}(-1)^{j+1}\det(A[j|1]) + \cdots + a_{in}(-1)^{j+n}\det(A[j|n]) \tag{16}$$

If $i = j$, this is the cofactor expansion along the i-th row. Suppose now that $i \neq j$. Let A' be the matrix A with the j-th row replaced by the i-th row. Then, by alternation, $\det(A') = 0$. Now, (16) is the cofactor expansion of A' along the j-th row, and therefore, it is 0. This proves that $A\mathrm{adj}(A) = \det(A)I$, and by definition of the adjoint, $\mathrm{adj}(A)A = A\mathrm{adj}(A)$, so we are done.

4.4 The multiplicativity of the determinant

The multiplicativity of the determinant is the property: $\det(AB) = \det(A)\det(B)$. This turns out to be a very strong property, from which all other properties follow readily in LAP.

Even the C-H theorem follows readily from the multiplicativity of det: from the multiplicativity of the determinant we have that $\det(I_{12}AI_{12}) = \det(I_1)\det(A)\det(I_1) = \det(A)$ for any matrix A. Suppose we want to prove the C-H theorem for some $n \times n$ matrix M. Define A as follows:

$$A = \begin{pmatrix} a & b & R \\ c & d & P \\ S & Q & M \end{pmatrix} = \begin{pmatrix} 0 & 0 & e_i^t \\ 0 & 0 & 0 \\ e_j & 0 & M \end{pmatrix}$$

Let $C_1C_2C_3\cdots C_{n+2}$ be the char poly of A (and $C_3\cdots C_{n+2}$ the char poly of M). From Berkowitz's algorithm it is easy to see that for A defined this way the bottom row of C_1C_2 is given by:

$$e_i^t M^n e_j \quad e_i^t M^{n-1} e_j \quad \ldots \quad e_i^t I e_j \quad 0$$

so the bottom row of $C_1C_2C_3\cdots C_{n+2}$ is simply $e_i^t p(M) e_j$ where p is the char poly of M.

On the other hand, using $\det(A) = \det(I_{12}AI_{12})$ and Berkowitz's algorithm, we have that:

$$\det(A) = \det \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & e_i^t \\ 0 & e_j & M \end{pmatrix} = 0$$

so that $e_i^t p(M) e_j = 0$, and since we can choose any i, j, we have that $p(M) = 0$.

What about the other direction? That is, can we prove the following implication in LAP:

C-H theorem $\Longrightarrow$ Multiplicativity of the determinant?

The answer is "yes," *if* LAP can prove the following:

$$\det(A) = 0 \rightarrow AB \neq I \tag{17}$$

That is, LAP can prove the multiplicativity of the determinant from the C-H theorem and (17). The proof of this is quite long, however, and the reader is once more directed to [8, Chapter 6].

We suspect, however, that LAP can prove (17) from the C-H theorem, so that the C-H Theorem is enough to prove multiplicativity. At this point, we do *not* have a LAP proof of (17) from the C-H theorem.

5 Propositional Translations

For a fixed effectively-presented underlying field K we can translate the theorems of LA and LAP into families of tautologies with short propositional proofs. These translations are a potential tool for proving independence results. For example, if we can prove that for $K = \mathbb{Z}_2$ the translation of $AB = I \rightarrow BA = I$ does not have polynomial size proofs in bounded-depth Frege with mod 2 gates, then it will follow that $AB = I \rightarrow BA = I$ is not a theorem of LA (over any field, since the theory LA is field independent).

Let α be a formula of LA or LAP, and let σ be an object assignment of natural numbers to all free index variables in α, and to all terms of the form $\mathsf{r}(A), \mathsf{c}(A)$. Let $|\sigma|$ be the largest value assigned by σ. To each variable of type field in α we assign one or more (depending on the underlying field) propositional variables (whose values determine a field element), and to each matrix variable A we assign enough propositional variables to determine all entries in A (where the size of A is determined by σ). Now α and σ translate into a propositional formula $||\alpha||_\sigma$ of size polynomial in $|\sigma|$ which is valid whenever α is true in the standard model under σ over the field K. The method of translation is similar to those described in Chapter 9 of [7].

If α is a theorem of LA then we show that $||\alpha||_\sigma$ has a Frege proof of size bounded by a polynomial in $|\sigma|$. In fact, we prove an even tighter result. When the underlying field is $\mathbb{Z}_p$, p a prime, the theorems of LA translate into propositional tautologies with short bounded-depth $\mathrm{AC}^0[p]$ proofs, where $\mathrm{AC}^0[p]$ allows $\mathsf{MOD}_{p,i}$ gates (i.e., modular gates, for counting modulo the prime p). We also show that the theorems of LAP have quasi-poly-bounded ($O(2^{\log^2 n})$) Frege proofs.

We point out in section 6.2 that the theorems of LAP have polynomial size Extended Frege proofs. For all the details see Chapter 7 of [8].

6 Proofs of the C-H theorem

The main result of this paper is a feasible proof of the Cayley-Hamilton (C-H) theorem. This result is important because it gives us a feasible proof of correctness of Berkowitz's algorithm, feasible proofs of hard matrix identities, and feasible proofs of the main principles of Matrix Algebra (specifically: axiomatic definition of the determinant, cofactor expansion formula, and multiplicativity of the determinant).

We believe that ours is the first feasible proof of the C-H theorem. The traditional proofs of the C-H theorem are infeasible, as they rely on the Lagrange expansion of the determinant, which for matrices of size n, has $n!$ terms of size n.

Our proof is formalized in the theory $\forall$LAP, which is obtained from LAP by introducing Π_1^M formulas.

Definition 6.1 We define Π_0^M to be the set of formulas over $\mathcal{L}_{\mathsf{LAP}}$ ("M" stands for matrix). We define Π_1^M to be the set of formulas in Π_0^M together with formulas of the form $(\forall A \leq n)\alpha$, where $\alpha \in \Pi_0^M$, and where $(\forall A \leq n)\alpha$ abbreviates:

$$(\forall A)((\mathsf{r}(A) \leq n \wedge \mathsf{c}(A) \leq n) \supset \alpha)$$

where A is a matrix variable, *not* contained in the index term n.

To form $\forall$LAP from LAP we add two LK-style rules $\forall$ $\leq$-left and $\forall$ $\leq$-right to $\forall$LAP, and augment the induction rule to allow induction over Π_1^M formulas. When the underlying field is finite or the rationals, this theory $\forall$LAP can be interpreted in the feasible theory $\mathbf{V}_1^1$. (See Section 5.5 of [7] for a definition of $\mathbf{V}_1^1$.)

The basic idea behind the proof 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 1×1 matrices, after at most $n =$ (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. The matrix B is obtained from A, by transposing the first row and column with the k-th row and column, respectively, and then deleting the first i rows and columns; finding k and i can be done in polytime.

It turns out that we do not need multiplicative inverses for field elements to prove the C-H theorem; that is, we do not need the function $^{-1}$. Berkowitz's algorithm does not compute inverses of field elements, and we do not need to reason with inverses in our proof of the C-H theorem. Thus, the C-H Theorem holds for commutative rings. On the other hand, we *do* use inverses in our proof of the multiplicativity of the determinant. It is an interesting question whether it is possible to give a feasible proof of the multiplicativity of the determinant for commutative rings.

Since LAP proves the equivalence of the C-H theorem, the axiomatic definition of the determinant, and the cofactor expansion (see Theorem 4.2), we conclude that these principles have feasible proofs as well. Using Gaussian Elimination, and the feasible proof of the C-H theorem, we also give a feasible proof of the multiplicativity of the determinant.

6.1 A feasible proof of the C-H theorem

In this section we give some more details of the feasible proof of the C-H theorem. A complete exposition is given in [8, Chapter 8].

We define the system $\forall$LAP to be similar to LAP, but we allow Π_1^M formulas. The underlying logic is again based on Gentzen's sequent system LK. Whereas LAP needs only the propositional rules of LK, we now need the rules for introducing a universal quantifier on the left and on the right of a sequent:

$$\begin{array}{rc} \text{left} & \dfrac{\mathsf{r}(T)\le n, \mathsf{c}(T)\le n, \alpha(T), \Gamma \to \Delta}{(\forall X \le n)\alpha(X), \Gamma \to \Delta} \\ \text{right} & \dfrac{\mathsf{r}(A)\le n, \mathsf{c}(A)\le n, \Gamma \to \Delta, \alpha(A)}{\Gamma \to \Delta, (\forall X \le n)\alpha(X)} \end{array}$$

where T is any term of type matrix, and n is any term of type index. Also, in $\forall$-introduction-right, A is a variable of type matrix that does not occur in the lower sequent, **and** in both rules α is a Π_0^M formula, because we just want a single matrix quantifier.

Note that $\forall$LAP still has the induction rule (3), and hence allows induction over Π_1^M formulas.

Theorem 6.1 $\forall$LAP proves the C-H theorem.

Proof. We prove that for all $n \times n$ matrices A, $p_A(A) = 0$, by induction on n. The **Basis Case** is trivial: if $A = (a_{11})$, then the char poly of A is $x - a_{11}$. We use the following strong induction hypothesis: $(\forall A \le n)p_A(A) = 0$. Thus, in our **Induction Step** we prove:

$$(\forall M \le n)p_M(M) = 0 \to (\forall A \le n+1)p_A(A) = 0 \quad (18)$$

So let A be an $(n+1) \times (n+1)$ matrix, and assume that we have $(\forall M \le n)p_M(M) = 0$. Then, by the results in Section 4, we have that for all $1 \le i < j \le n-1$, $p_{(I_{ij}AI_{ij})} = p_A$.

Suppose now that the i-th row (column) of $p_A(A)$ is not zero. Then, the first row (column) of $I_{1i}p_A(A)I_{1i}$ is not zero. But:

$$I_{1i}p_A(A)I_{1i} = p_A(I_{1i}AI_{1i}) = p_{(I_{1i}AI_{1i})}(I_{1i}AI_{1i})$$

and the first row and column of $p_{(I_{1i}AI_{1i})}(I_{1i}AI_{1i})$ are zero by Lemma 6.1 below (letting $C = I_{1i}AI_{1i}$). Thus, contradiction; it follows that $p_A(A) = 0$. This argument can be clearly formalized in $\forall$LAP. □

Lemma 6.1 LAP proves that if $p_{C[1|1]}(C[1|1]) = 0$, then the first row and the first column of $p_C(C)$ are zero.

Proof. We restate the Lemma using the usual notation of A and $M = A[1|1]$. Thus, we want to show that LAP proves the following: if $p_M(M) = 0$, then the first row and the first column of $p_A(A)$ are zero. For clarity we let $p = p_A$ and $q = p_M$.

The proof is by induction on the size of M. The **Basis Case** is when M is a 1×1 matrix. Let p_2, p_1, p_0 be the coefficients of the char poly of A, and let q_1, q_0 be the coefficients of the char poly of M. By assumption $q_1M + q_0I = 0$. Note that I is also a 1×1 matrix. From Berkowitz's algorithm we know that:

$$\begin{aligned} \begin{pmatrix} p_2 \\ p_1 \\ p_0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ -a_{11} & 1 \\ -RS & -a_{11} \end{pmatrix} \begin{pmatrix} q_1 \\ q_0 \end{pmatrix} \\ &= \begin{pmatrix} q_1 \\ -a_{11}q_1 + q_0 \\ -RSq_1 - a_{11}q_0 \end{pmatrix} \end{aligned} \quad (19)$$

Note that:

$$A^2 = \begin{pmatrix} a_{11}^2 + RS & a_{11}R + RM \\ a_{11}S + MS & SR + M^2 \end{pmatrix}$$

We must now show that the first row and column of $p_A(A) = p_2A^2 + p_1A + p_0I$ are zero. We just show that the $(1,2)$ entry is zero; the rest follow just as easily. From (19) we see that the $(1,2)$ entry of $p_A(A)$ is given by:

$$\begin{aligned} &(a_{11}R + RM)q_1 + R(-a_{11}q_1 + q_0) + 0(-RSq_1 - a_{11}q_0) \\ &= R(Mq_1 + q_0) = 0 \end{aligned}$$

Note that it is actually possible, in the Basis Case, to show that $p_A(A) = 0$ (as this is *true*), not just the first row and column of $p_A(A)$. However, this seems infeasible to carry out in the Induction Step.

We prove the **Induction Step** with three claims. We assume that M is an $(n-1) \times (n-1)$ matrix, where $n-1 \ge 1$. We let $p = p_A$ and $q = p_M$, that is, p, q are the char polys of $A, M = A[1|1]$, respectively. Define w_k, X_k, Y_k, Z_k as follows:

$$\begin{aligned} A &= \begin{pmatrix} w_1 & X_1 \\ Y_1 & Z_1 \end{pmatrix} = \begin{pmatrix} a_{11} & R \\ S & M \end{pmatrix} \\ A^{k+1} &= \begin{pmatrix} w_{k+1} & X_{k+1} \\ Y_{k+1} & Z_{k+1} \end{pmatrix} \\ &= \begin{pmatrix} w_k & X_k \\ Y_k & Z_k \end{pmatrix} \begin{pmatrix} a_{11} & R \\ S & M \end{pmatrix} \quad \text{for } k \ge 1 \end{aligned}$$

Note that w_k, X_k, Y_k, Z_k cannot be defined in LAP as we cannot define new matrices recursively. However, all that we need in the following proof are entries of powers of A, which can be expressed in LAP. The entry w_k, and the submatrices X_k, Y_k, Z_k are there to make the proof more human readable; for example, instead of w_k we could write $\mathsf{e}(\mathsf{P}(k,A),1,1)$, or instead of X_k we could write $\lambda ij\langle 1, n-1, \mathsf{e}(\mathsf{P}(k,A),1,j+1)\rangle$, but then the proof would be difficult to read.

It is easy to see that LAP proves the following equations:

$$\begin{aligned} w_{k+1} &= a_{11}w_k + X_kS \\ X_{k+1} &= w_kR + X_kM \\ Y_{k+1} &= a_{11}Y_k + Z_kS \\ Z_{k+1} &= Y_kR + Z_kM \end{aligned} \tag{20}$$

As was mentioned above, we are going to prove that the first row and column consist of zeros with Claims 6.1, 6.2, and 6.3. Claim 6.3 follows from Claim 6.2 using the fact that A and A^t have the same char poly (the details are provided in the proof of Claim 6.3). For the other two claims we are going to put $p_A(A)$ in a special form. Using Berkowitz's algorithm, it is easy to show in LAP that:

$$p(A) = (A - a_{11}I)q(A) - \sum_{k=1}^{n-1} q_k \sum_{i=0}^{k-1}(RM^iS)A \tag{21}$$

and thus, to show that the first column of $p(A)$ is zero, it is enough to show that the first columns of $(A - a_{11}I)q(A)$ and $\sum_{k=1}^{n-1} q_k \sum_{i=0}^{k-1}(RM^iS)A$ are the same. This is the strategy for proving Claims 6.1 and 6.2.

Claim 6.1 The upper-left entry of $p(A)$ is zero.

Proof. Using (20) we obtain:

$$\begin{cases} w_0 = 1 \\ w_1 = a_{11} \\ w_{k+1} = a_{11}w_k + \sum_{i=0}^{k-1}(RM^iS)w_{k-1-i} \end{cases} \tag{22}$$

For $k \geq 1$. The top left entry of $(A - a_{11}I)q(A)$ is given by

$$\sum_{k=1}^{n-1} q_k(w_{k+1} - a_{11}w_k) \tag{23}$$

(notice that we can ignore the term $k = 0$ since the top left entry of A is the same as the top left entry of $a_{11}I$). We can compute $(w_{k+1} - a_{11}w_k)$ using the recursive definitions of w_k (given by (22) above):

$$\begin{aligned} w_{k+1} - a_{11}w_k &= a_{11}w_k + \sum_{i=0}^{k-1}(RM^iS)w_{k-1-i} - a_{11}w_k \\ &= \sum_{i=0}^{k-1}(RM^iS)w_{k-1-i} \end{aligned}$$

Thus, (23) is equal to

$$\sum_{k=1}^{n-1} q_k \sum_{i=0}^{k-1}(RM^iS)w_{k-1-i}$$

This proves that the top left entry of $p(A)$ is zero (see equation (21) and the explanation below it). □

Claim 6.2 The $(n-1) \times 1$ lower-left submatrix of $p(A)$ is zero.

Proof. Using (20) we obtain ($k \geq 1$):

$$\begin{cases} Y_0 = 0 \\ Y_1 = S \\ Y_{k+1} = a_{11}Y_k + (M^kS) + \sum_{i=0}^{k-2}(RM^iS)Y_{k-1-i} \end{cases} \tag{24}$$

The lower-left $(n-1) \times 1$ submatrix of $(A - a_{11}I)q(A)$ is given by

$$\sum_{k=0}^{n-1} q_k(Y_{k+1} - a_{11}Y_k)$$

and by (24) we have that for $k \geq 2$, $Y_{k+1} - a_{11}Y_k$ is given by:

$$\begin{aligned} &= \left(a_{11}Y_k + M^kS + \sum_{i=0}^{k-2}(RM^iS)Y_{k-1-i}\right) - a_{11}Y_k \\ &= M^kS + \sum_{i=0}^{k-2}(RM^iS)Y_{k-1-i} \end{aligned}$$

ans, therefore, $\sum_{k=0}^{n-1} q_k(Y_{k+1} - a_{11}Y_k)$ is given by:

$$\begin{aligned} &= q_0(Y_1 - a_{11}Y_0) + q_1(Y_2 - a_{11}Y_1) \\ &\quad + \sum_{k=2}^{n-1} q_k \left(M^kS + \sum_{i=0}^{k-2}(RM^iS)Y_{k-1-i}\right) \\ &= q(M)S + \sum_{k=2}^{n-1} q_k \sum_{i=0}^{k-2}(RM^iS)Y_{k-1-i} \end{aligned}$$

and by the IH, $\sum_{k=0}^{n-1} M^kS = q(M)S = 0$, and by definition $Y_0 = 0$, thus we can conclude that:

$$\sum_{k=0}^{n-1} q_k(Y_{k+1} - a_{11}Y_k) = \sum_{k=1}^{n-1} q_k \sum_{i=0}^{k-1}(RM^iS)Y_{k-1-i}$$

But the RHS of the above equation is equal to the $(n-1) \times 1$ lower-left submatrix of $\sum_{k=1}^{n-1} q_k \sum_{i=0}^{k-1}(RM^iS)A^{k-1-i}$, and hence the claim follows (once again, see equation (21) and the explanation below it). □

Claim 6.3 The $1 \times (n-1)$ upper-right submatrix of $p(A)$ is zero.

Proof. To prove this claim we use the fact that $p_A = p_{A^t}$ and Claim 6.2. The crucial observation is that the $(n-1) \times 1$ lower-left submatrix of $(A^t)^k$ is X_k^t. Now, we know that p is also the char polynomial of A^t, so by Claim 6.2, we know that the $(n-1) \times 1$ lower-left submatrix of $p(A^t)$ is zero. Thus the $(n-1) \times 1$ lower-left submatrix of $(p(A))^t$ is zero, and therefore the $1 \times (n-1)$ upper-right submatrix of $p(A)$ is zero, and hence the claim follows. □ □

6.2 Other propositional proofs

We also show how to translate feasible proofs of the C-H theorem over $\mathbb{Z}_2$ into families of NC^2-Frege (quasi-poly bounded Frege) proofs with the permutation rule, and also into poly bounded Frege with propositional quantifiers. Permutation Frege [1, 9] is a fragment of Substitution Frege, which corresponds to reasoning with poly-time concepts. The fragment of Quantified Frege that we use is tree-like and all formulas only need one block of universal quantifiers, and this can be p-simulated by Extended Frege.

6.3 Gaussian Elimination

The correctness condition of Gaussian Elimination states that after performing the algorithm, the resulting matrix is in row-echelon form. This condition of correctness can be expressed with a family of tautologies (depending on the underlying field, as usual), of size polynomial in the size of the matrix. We outline how to prove these tautologies with uniform polysize eFrege proofs. This gives us feasible proofs of correctness of Gaussian Elimination.

This result is interesting because we do not know how to give a proof of correctness of Berkowitz's algorithm in its own complexity class. In other words, we do not know if we can prove the C-H theorem using NC^2 concepts, rather than (feasible) polynomial-time concepts.

We use the proof of correctness of Gaussian Elimination to give a direct feasible proof (as opposed to a proof via the feasible proof of the C-H theorem) of $AB = I \rightarrow BA = I$.

7 Conclusion and Open Problems

At this point, it is not known if there are poly-bounded Frege proofs, or even quasi-poly-bounded Frege proofs of hard matrix identities or of the Cayley-Hamilton (C-H) theorem. To repeat using the language of circuit complexity: we know that hard matrix identities, as well as the C-H theorem, have poly-bounded P/poly-Frege proofs, but it is not known if they have poly-bounded NC^i-Frege proofs, for any i. Since Berkowitz's algorithm is an NC^2 algorithm, it is tempting to conjecture that they all have NC^2-Frege proofs.

Here are some other open questions. More details can be found in Chapter 9 of [8].

1. Show that LA cannot prove $AB = I \rightarrow BA = I$. The most obvious approach is to construct a model $\mathcal{M}$ of LA such that $\mathcal{M} \not\models AB = I \rightarrow BA = I$. Alasdair Urquhart (private communication) suggested another approach as follows: He showed that if LA $\vdash AB = I \rightarrow BA = I$ then the Propositional Pigeonhole Principal has polynomial size bounded-depth Frege proofs with mod 2 gates. The latter is believed to be unlikely.

2. Is $AB = I \rightarrow BA = I$ "Complete"? Theorem 4.1 states that LAP proves that the C-H theorem implies $AB = I \rightarrow BA = I$. Could it be that LAP+ C-H is a conservative extension of LA + $AB = I \rightarrow BA = I$?

3. Does LAP prove det$(A) = 0 \rightarrow AB \neq I$? If so, then LAP proves the equivalence of the multiplicativity of the determinant with the other three principles of Section 4.

Acknowledgments: Our thanks to Sam Buss for fruitful comments resulting from the careful reading of the source of this paper: the first author's PhD thesis [8].

References

[1] N. H. Arai. Tractability of cut-free Gentzen type propositional calculus with permutation inference. 1995.

[2] S. J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18(3):147–150, 1984.

[3] M. Bonet, S. Buss, and T. Pitassi. Are there hard examples for frege systems? *Feasible Mathematics*, II:30–56, 1994.

[4] S. R. Buss. *Bounded Arithmetic*. Studies in proof theory. Napoli, 1986.

[5] S. A. Cook. Feasibly constructive proofs and the propositional calculus. *Proc. 7th ACM Symposium on the Theory of Computation*, pages 83–97, 1975.

[6] S. A. Cook. A taxonomy of problems with fast parallel algorithms. *Information and Computation*, 64(13):2–22, 1985.

[7] J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Cambridge, 1995.

[8] M. Soltys. *The Complexity of Derivations of Matrix Identities*. PhD thesis, University of Toronto, Department of Mathematics, 2001. Available from the ECCC server (under Theses).

[9] A. Urquhart. The symmetry rule in propositional logic. *Discrete Applied Mathematics*, 96–97:177–193, 1998.

[10] J. von zur Gathen. Parallel linear algebra. In J. H. Reif, editor, *Synthesis of Parallel Algorithms*, pages 574–617. Morgan and Kaufman, 1993.