The proof theoretic strength of the Steinitz Exchange Theorem

Michael Soltys

May 3, 2006

Abstract

We show that the logical theory **QLA** proves the Cayley-Hamilton theorem from the Steinitz Exchange theorem together with a strengthening of the linear independence principle. Since **QLA** is a fairly weak theory (in the sense that its quantifier-free fragment, **LA**, translates into tautologies with $\mathbf{TC}^0$-Frege proofs—when restricted to the field $\mathbb{Q}$ of the rationals), it follows that the proof complexity of matrix algebra can be distilled to the Steinitz Exchange theorem.

1 Introduction

The theory **LA** is a field-independent logical theory for matrix algebra. **LA** proves all the ring properties of matrices (e.g., $A(BC) = (AB)C$) and restricted to the field $\mathbb{Q}$ of the rationals translates in $\mathbf{TC}^0$-Frege. **QLA** is a quantified theory axiomatized by all the theorems of **LA**. As these theories are described in detail elsewhere ([5]), suffice it to say that **LA** is a natural theory for matrix algebra, with three sorts: indices, field elements, and matrices, and axiomatized by the usual number theoretic axioms for indices, the usual field axioms, and induction over open formulas. See the appendix for a table giving a brief description of **LA**.

In practice, **QLA** is just strong enough to prove all the ring properties of matrices, and it is conjectured that it is too weak to prove properties of matrix inverses. We show in this paper that inverse properties of matrices, of which the Cayley-Hamilton theorem is the example *par excellence* (because it gives a (non-zero) annihilating polynomial for a matrix, and effectively shows that the adjoint is the inverse), can be shown in **QLA** from the Steinitz Exchange theorem and a strengthening of the linear independence principle. A preliminary version of this paper, containing some of the results mentioned here, appeared in [4]. For the algebra material in this paper the reader can consult, for example, [1] or [3].

Let the **Strong Linear Independence (SLI)** principle be the following assertion: if $\{v_1, \ldots, v_m\}$ are $n \times 1$, non-zero, linearly dependent vectors, then there exists a $1 \leq k < m$ such that $\{v_1, \ldots, v_k\}$ is linearly independent, but

$\{v_1, \ldots, v_{k+1}\}$ is linearly dependent. This can be stated easily as a **QLA**-formula by encoding the vectors as the columns of a matrix, a practice we follow implicitly throughout the paper.

Recall that the **Steinitz Exchange Theorem (SET)** states the following: if T is a total set (i.e., it spans the whole vector space), and E is a linearly independent set, then there exists a subset $F \subseteq T$, such that $|F| = |E|$, and $(T - F) \cup E$ is total. Note that totality can be stated as follows: $\exists X[TX = I]$, and also that the standard proof of **SET** (as, for example, given in [2, pg. 216]) can be formalized in $\exists$**LA** (so it has a *polytime* proof).

Note that both the number k in **SLI** and the set F in **SET** can be computed with $\mathbf{NC}^2$ algorithms (polysize circuits of depth $O(\log^2 n)$). To compute the k, we compute the rank of $\{v_1, \ldots, v_i\}$ and $\{v_1, \ldots, v_{i+1}\}$ independently for all $i < m$, and we let $k = i$ for the first i for which the two sets have the same rank. To compute the F in **SET** we list all the vectors in E followed by all the vectors in T, and for each $i < |E|+|T|$ we check if the first i vectors and the first $(i+1)$ vectors have the same rank; if they do, we put the $(i+1)$ vector in F, and we stop when $|F| = |E|$. As rank can be computed in $\mathbf{NC}^2$ with Mulmuley's algorithm, the claim follows (for an exposition of Mulmuley's algorithm, see [6]).

Finally, note once again that both **SET** and **SLI** can be shown in the theory $\exists$**LA** using a straightforward induction over Σ_1^B formulas (i.e., formulas of the form $(\exists X \leq n)\alpha$, where X is a matrix of size at most n, and α has not matrix quantifiers).

Csanky's algorithm for computing the characteristic polynomial of a matrix uses Newton's symmetric polynomials, which are defined as follows: $s_0 = 1$, and for $1 \leq k \leq n$,

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

Then, $p_A(x) := s_0 x^n - s_1 x^{n-1} + s_2 x^{n-2} - \cdots \pm s_n x^0$. Note that we can express the coefficients of the characteristic polynomial (computed by Csanky's algorithm) as a term in the language of **LAP** (which is **LA** together with the matrix powering function), as the next lemma shows. Note that we require fields of characteristic 0 to run Csanky's algorithm, and so the underlying assumption throughout this paper is that the field is $\mathbb{Q}$. (See [6] for a more detailed exposition.)

Lemma 1 *p_A can be given as a term of* **LAP**.

Proof. We restate (1) in matrix form: $s = Ts - b$ where s, T, b are given, respectively, as follows:

$$\begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & \dots \\ \frac{1}{2}\mathrm{tr}(A) & 0 & 0 & \dots \\ \frac{1}{3}\mathrm{tr}(A^2) & \frac{1}{3}\mathrm{tr}(A) & 0 & \dots \\ \frac{1}{4}\mathrm{tr}(A^3) & \frac{1}{4}\mathrm{tr}(A^2) & \frac{1}{4}\mathrm{tr}(A) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \begin{pmatrix} \mathrm{tr}(A) \\ \frac{1}{2}\mathrm{tr}(A^2) \\ \vdots \\ \frac{1}{n}\mathrm{tr}(A^n) \end{pmatrix}$$

Then $s = -b(I - T)^{-1}$. Note that $(I - T)$ is an invertible matrix as it is lower triangular, with 1s on the main diagonal, and so its inverse can be computed easily, and shown that it is computed correctly, in **LAP**. □

Note that we do not have matrix powering in **QLA**, but we show in lemma 3 that we can prove the existence of the powers of matrices in **QLA** from **SET**, and thus **QLA** with **SET** can "simulate" **LAP**.

2 A proof of the Cayley-Hamilton Theorem

Let A be an $n \times n$ matrix, and $p_A(x)$ its characteristic polynomial computed by Csanky's algorithm. By the comment at the end of the preceding section, we can prove the existence of $p_A(x)$ in **QLA** with **SET**. The **CHT** states that $p_A(A) = 0$. To show that, we are going to prove in **QLA** that for all i, $p_A(A)e_i = 0$, where $B_0 = \{e_1, \dots, e_n\}$ is the standard basis.

Theorem 1 *The theory* **QLA** *proves from* **SLI** *and* **SET** *that for all the vectors e_i in the standard basis, $(p_A(A))e_i = 0$, where p_A is the characteristic polynomial of the matrix A as computed by Csanky's algorithm. It follows that $p_A(A) = 0$, i.e., the Cayley-Hamilton theorem.*

Proof. Consider the set $W = \{e_i, Ae_i, \dots, A^n e_i\}$. By the principle of linear independence ($(n+1)$ n-vectors must be linearly dependent, which follows from **SET** in **QLA** by lemma 2 below) we know that W must be linearly dependent. By **SLI** there exists a $k \leq n$ such that the first k vectors of W are linearly independent, but the first $(k+1)$ vectors are not; let $W_0 = \{e_i, Ae_i, \dots, A^{k-1}e_i\}$ where k is the largest power of A such that W_0 is linearly independent.

As an aside, note that showing the existence of such a k seems to require induction over Σ_1^B formulas, which we do not have in **QLA**. On the other hand, any subset of a linearly independent set is linearly independent, and any superset of a linearly dependent set is linearly dependent (and furthermore, this can be shown in **QLA**). It thus follows that we can find this k with binary search, where we ask if $\{e_i, Ae_i, \dots, A^j e_i\}$ is linearly independent or not, and eliminate half of the remaining candidates for k in one step.

Then $A^k e_i$ 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 $g(x) = x^k + c_1 x^{k-1} + \cdots + c_k$, then $g(A)e_i = 0$.

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

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

We show in lemma 4 below that **LAP** proves that $p_{A_g} = g$, and conclude that $(p_{A_g}(A))e_i = 0$. (Remember that by lemma 3 we can simulate **LAP** in **QLA** with **SET**.)

Now, we extend W_0 to $B = W_0 \cup \{e_{j_1}, \dots, e_{j_{n-k}}\}$, a basis, using **SET** (let $T = B_0$, the standard basis, and let $E = W_0$, which is linearly independent, so $B = (T - F) \cup E$ for some F). Let M_B be the matrix whose columns are the vectors of B, and so M_B is invertible (because the columns of M_B are linearly independent, and so by lemma 2 M_B has a right-inverse, and hence a two-sided-inverse) and

$$A \cdot M_B = M_B \cdot \begin{pmatrix} A_g & E_1 \\ 0 & E_2 \end{pmatrix}$$

where $E = \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}$ is defined by $E = M_B^{-1} A'$ where A' is composed of columns $j_1, \dots, j_{n-k}$ of A (note that M_B is a *change of basis*). Thus:

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

i.e., they are similar matrices.

LAP proves that if $C_1 \sim C_2$ then $p_{C_1}(x) = p_{C_2}(x)$, i.e. that similar matrices have the same characteristic polynomial (this follows easily from the fact that $\text{tr}(A) = \text{tr}(PAP^{-1})$—since $\text{tr}(AB) = \text{tr}(BA)$—and (1)), and we show that **LAP** proves that if

$$C = \begin{pmatrix} C_1 & * \\ 0 & C_2 \end{pmatrix} \tag{3}$$

then $p_C(x) = p_{C_1}(x) \cdot p_{C_2}(x)$ (see lemma 5 below for a proof of this; note that the lemma proves this result for a transpose of (3), but since $\text{tr}(A) = \text{tr}(A^t)$, we have the result for (3) as well).

We conclude that $p_A(A)e_i = (p_{A_g}(A) \cdot p_E(A))e_i = p_E(A) \cdot (p_{A_g}(A)e_i) = 0$. Since this holds for all the vectors e_i in the standard basis, it follows that $p_A(A) = 0$. □

Lemma 2 **QLA** *proves that the following principles are consequences of* **SET***:*

1. *$(\exists B \neq 0)[AB = I \vee AB = 0]$,*
2. *Linear Independence ($n+1$ vectors in $\mathbb{F}^n$ are linearly independent),*
3. *Every matrix has an annihilating polynomial,*
4. *$AB = I \supset BA = I$.*

Note that the understanding is that the matrices A, B mentioned above are square *matrices.*

Proof. Let A be an $n \times n$ matrix, and suppose that $AB \neq 0$ for any $B \neq 0$. Then the columns of A are linearly independent, and so by **SET** they are total (just take $T = B_0$, the standard basis), and in particular there exist B_i's such that $AB_i = e_i$, so $A[B_1 B_2 \ldots B_n] = I$.

1 implies 2: suppose that we have $n+1$ vectors, and arrange them as the columns of a matrix A (and so A is an $n \times (n+1)$ matrix). If $AB \neq 0$ for all $B \neq 0$, then the same holds for A' which is A with a row of zeros appended (and so A' is a *square* matrix), and hence there exists $B' \neq 0$ such that $A'B' = I$ which is not possible.

2 implies 3: consider $I, A, A^2, \ldots, A^{n^2}$. This set is linearly dependent (as vectors), and so there exists $C \neq 0$ such that $c_0 I + c_1 A + \cdots c_{n^2} A^{n^2} = 0$ giving us the coefficients of an annihilating polynomial of A.

3 implies 4: suppose that $AB = I$, and let p be an annihilating polynomial of A. If $p_0 \neq 0$, i.e., the constant coefficient of p is not zero, then from $p(A) = 0$ we can obtain a two sided inverse of A. If $p_0 = 0$, then let q be a polynomial with $q_0 \neq 0$ such that $p(A) = q(A)A^s$. Since $AB = I$, $A^s B^s = I$, so $q(A) = 0$, so once again we have a two sided inverse. Since $AB = I$ implies (in **LA**) that $A(BA - I) = 0$, we know that $BA = I$. Note that this argument requires finding an s such that $p(A) = q(A)A^s$, which effectively means finding the first non-zero coefficient of p. Since p is represented as a vector of coefficients, i.e., $p = [p_m p_{m-1} \ldots p_0]^t \neq 0$, this means finding the least s such that $p_s \neq 0$, which can be done with **LA**-induction. □

Lemma 3 **QLA** *can prove the existence of powers of a matrix from* **SET**.

Proof. Let $\mathrm{POW}(A, n)$ be the formula:

$$\exists \langle X_0 X_1 \ldots X_n \rangle (\forall i \leq n)[X_0 = I \wedge (i < n \supset X_{i+1} = X_i * A)] \tag{4}$$

The size of $\langle X_0 X_1 \ldots X_n \rangle$ can be bounded as it is a $r(A) \times (r(A) \cdot (n+1))$ matrix.

We now show that $\mathbf{QLA} \vdash (\exists B \neq 0)[AB = I \vee AB = 0] \supset \mathrm{POW}(A, n)$. Note that in lemma 2 we showed that the LHS of this implication is provable in **QLA** from **SET**.

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. Then $N^n = 0$, and $(I-N)^{-1} = I + N + N^2 + \ldots + N^{n-1} =$

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

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 $CB = I$.

Next, show that $B = I + N + N^2 + \cdots + N^{n-1}$, again, by induction on the rows of B, starting with the bottom row. Thus, B contains $I, A, A^2, \ldots, A^{n-1}$ in its top rows, and $\mathrm{POW}(A, n)$ follows. □

Lemma 4 *Let A_g be the companion matrix of the polynomial g (see (2)). Then* **QLAP** *shows that the characteristic polynomial of A_g is precisely g, i.e., $p_{A_g} = g$, and furthermore* **QLAP** *shows that $p_{A_g}(A_g) = 0$, i.e., we have the* **CHT** *for companion matrices.*

Proof. The following proof is not *complex*, but it is *complicated*; it can be formalized in **LAP** (and hence in **QLA** with **SET**), but it is quite technical. The matrix $A := A_g$ (we drop the subscript for readability) is a $k \times k$ matrix, with 1s below the main diagonal, and zeros everywhere else except (possibly) in the last column where it has the negations of the coefficients of $g(x)$ (again, see (2) for a definition of A_g).

Divided A into four quadrants, with the upper-left containing just 0. Let $R = \begin{pmatrix} 0 & \ldots & 0 & -c_k \end{pmatrix}$ be the $1 \times (k-1)$ row vector in the upper-right quadrant, i.e., the first row of A without the first entry. Let $S = e_1$ be the $(k-1) \times 1$ column vector in the lower-left quadrant, i.e., the first column of A without the top entry. Finally, let M be the principal submatrix of A, $M = A[1|1]$; the lower-right quadrant.

Let $s_0, s_1, \ldots, s_k$ be the Newton's symmetric polynomials of A, as given by (1). To prove that $g(x) = p_A(x)$ we prove something stronger: we show that (i) for all $0 \leq i \leq k$ $(-1)^i s_i = c_i$, and (ii) $p_A(A) = 0$.

We show this by induction on the size of the matrix A. Since the principal submatrix of A (i.e., M) is *also* a companion matrix, we assume that for $i < k$, the coefficients of the symmetric polynomial of M are equal to the c_i's, and that $p_M(M) = 0$. (Note that the Basis Case of the induction is a 1×1 matrix, and it is trivial to prove.)

Since for $i < k$, $\mathrm{tr}(A^i) = \mathrm{tr}(M^i)$, it follows from (1) and the induction hypothesis that for $i < k$, $(-1)^i s_i = c_i$ (note that $s_0 = c_0 = 1$).

Next we show that $(-1)^k s_k = c_k$. By definition (i.e., by (1)) we have that s_k is equal to:

$$\frac{1}{k}(s_{k-1}\mathrm{tr}(A) - s_{k-2}\mathrm{tr}(A^2) + \cdots + (-1)^{k-2}s_1\mathrm{tr}(A^{k-1}) + (-1)^{k-1}s_0\mathrm{tr}(A^k)),$$

and by the induction hypothesis and the fact that for $i < k$ $\mathrm{tr}(A^i) = \mathrm{tr}(M^i)$ we have that this is equal to

$$\frac{1}{k}(-1)^{k-1}(c_{k-1}\mathrm{tr}(M) + c_{k-2}\mathrm{tr}(M^2) + \cdots + c_1\mathrm{tr}(M^{k-1}) + c_0\mathrm{tr}(A^k)),$$

and note that $\mathrm{tr}(A^k) = -kc_k + \mathrm{tr}(M^k)$, so this equals

$$\begin{aligned}&\frac{1}{k}(-1)^{k-1}\left[c_{k-1}\mathrm{tr}(M) + c_{k-2}\mathrm{tr}(M^2) + \cdots + c_1\mathrm{tr}(M^{k-1}) + c_0\mathrm{tr}(M^k)\right]\\ &\qquad + (-1)^k c_k.\end{aligned}$$

Observe that

$$\mathrm{tr}(c_{k-1}M + c_{k-2}M^2 + \cdots + c_1M^{k-1} + c_0M^k) = \mathrm{tr}(p_M(M)M) = \mathrm{tr}(0) = 0$$

since $p_M(M) = 0$ by the induction hypothesis. Therefore, $s_k = (-1)^k c_k$.

It remains to prove that $p_A(A) = \sum_{i=0}^k c_iA^{k-i} = 0$. First, show that for $1 \leq i \leq (k-1)$:

$$A^{i+1} = \left(\begin{array}{c|c} 0 & RM^i \\ \hline M^iS & \sum_{j=0}^{i-1} M^jSRM^{(i-1)-j} + M^{i+1} \end{array}\right) \tag{5}$$

(for A of the form given by (2), and R, S, M defined as in the first paragraph of the proof). Define w_i, X_i, Y_i, Z_i as follows:

$$\begin{aligned} A^{i+1} = \begin{pmatrix} w_{i+1} & X_{i+1} \\ Y_{i+1} & Z_{i+1} \end{pmatrix} &= \begin{pmatrix} w_i & X_i \\ Y_i & Z_i \end{pmatrix}\begin{pmatrix} 0 & R \\ S & M \end{pmatrix} \\ &= \begin{pmatrix} X_iS & w_iR + X_iM \\ Z_iS & Y_iR + Z_iM \end{pmatrix}\end{aligned} \tag{6}$$

We want to show that the right-most matrix of (6) is equal to the right-hand side of (5). First note that:

$$X_{i+1} = \sum_{j=0}^{i} w_{i-j}RM^j \qquad w_{i+1} = \sum_{j=0}^{i-1}(RM^jS)w_{i-1-j} \tag{7}$$

With the convention that $w_0 = 1$. Since $w_1 = 0$, a straight-forward induction shows that $w_{i+1} = 0$. Therefore, at this point the right-most matrix of (6) can be simplified to:

$$\begin{pmatrix} 0 & RM^i \\ Z_iS & Y_iR + Z_iM \end{pmatrix}$$

We have:

$$Y_{i+1} = M^iS + \sum_{j=0}^{i-2}(RM^jS)Y_{i-1-j} \qquad Z_{i+1} = M^{i+1} + \sum_{j=0}^{i-1} Y_{i-1-j}RM^j$$

By the same reasoning as above, $\sum_{j=0}^{i-2}(RM^jS)Y_{i-1-j} = 0$, so putting it all together we obtain the right-hand side of (5).

Using the induction hypothesis ($p_M(M) = 0$) it is easy to show that the first row and column of $p_A(A)$ are zero. Also, by the induction hypothesis, the term M^{i+1} in the principal submatrix of $p_A(A)$ disappears but leaves c_kI. Therefore, it will follow that $p_A(A) = 0$ if we show that

$$\sum_{i=2}^{k} c_{k-i} \sum_{j=0}^{i-2} M^j SRM^{(i-2)-j} \tag{8}$$

is equal to $-c_kI$.

Some observations about (8): for $0 \leq j \leq i-2 \leq k-2$, the first column of M^j is just e_{j+1}. And SR is a matrix of zeros, with $-c_k$ in the upper-right corner. Thus M^jSR is a matrix of zeros except for the last column which is $-c_ke_{j+1}$. Thus, $M^jSRM^{(i-2)-j}$ is a matrix with zeros everywhere, except in row $(j+1)$ where it has the bottom row of $M^{(i-2)-j}$ multiplied by $-c_k$. Let $\mathbf{m}^{(i-2)-j}$ denote the $1 \times (k-1)$ row vector consisting of the bottom row of $M^{(i-2)-j}$. Therefore, (8) is equal to:

$$-c_k \cdot \left(\begin{array}{c} \sum_{i=2}^{k} c_{k-i}\mathbf{m}^{(i-2)} \\ \hline \sum_{i=3}^{k} c_{k-i}\mathbf{m}^{(i-3)} \\ \hline \vdots \\ \hline \sum_{i=k}^{k} c_{k-i}\mathbf{m}^{(i-k)} \end{array} \right) \tag{9}$$

We want to show that (9) is equal to $-c_kI$ to finish the proof of $p_A(A) = 0$. To accomplish this, let l denote the l-th row of the matrix in (9) starting with the bottom row. We want to show, by induction on l, that the l-th row is equal to e_{k-l}.

The Basis Case is $l = 0$:

$$\sum_{i=k}^{k} c_{k-i}\mathbf{m}^{(i-k)} = c_0\mathbf{m}^0 = e_{k-1}^t,$$

and we are done.

For the induction step, note that $\mathbf{m}^{l+1}$ is equal to $\mathbf{m}^l$ shifted to the left by one position, and with

$$\mathbf{m}^l \cdot \left(\begin{array}{cccc} -c_{k-1} & -c_{k-2} & \ldots & -c_1 \end{array} \right)^t \tag{10}$$

in the last position. (Here the "$\cdot$" denotes the dot product of the two vectors.) We introduce some more notation: let $\mathbf{r}_l$ denote the $k-l$ row of (9). Thus $\mathbf{r}_l$ is $1 \times (k-1)$ row vector. Let $\overleftarrow{\mathbf{r}}_l$ denote $\mathbf{r}_l$ shifted by one position to the left, and with a zero in the last position. This can be stated succinctly in **LAP** as follows:

$$\overleftarrow{\mathbf{r}}_l \overset{\text{def}}{=} \lambda ij\langle 1,(k-1),e(\mathbf{r}_l,1,i+1)\rangle.$$

Based on (9) and (10) we can see that:

$$\mathbf{r}_{l+1} = \overleftarrow{\mathbf{r}}_l + [\mathbf{r}_l \cdot (\begin{array}{cccc} -c_{k-1} & -c_{k-2} & \dots & -c_1 \end{array})^t]e^t_{k-1} + c_l\mathbf{m}^0.$$

Using the induction hypothesis: $\overleftarrow{\mathbf{r}}_l = e^t_{k-(l+1)}$, and

$$\mathbf{r}_l \cdot (\begin{array}{cccc} -c_{k-1} & -c_{k-2} & \dots & -c_1 \end{array})^t = e^t_{k-l} \cdot (\begin{array}{cccc} -c_{k-1} & -c_{k-2} & \dots & -c_1 \end{array})^t = -c_l$$

so $\mathbf{r}_{l+1} = e^t_{k-(l+1)} - c_l e^t_{k-1} + c_l e^t_{k-1} = e^t_{k-(l+1)}$ as desired. This finishes the proof of the fact that the matrix in (9) is the identity matrix, which in turn proves that (8) is equal to $-c_k I$, and this ends the proof of $p_A(A) = 0$, which finally finishes the main induction argument, and proves the lemma. □

Lemma 5 **LAP** *proves that if A is a matrix of the form:*

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \tag{11}$$

where B and D are square matrices (not necessarily of the same size), and the upper-right corner is zero, then $p_A(x) = p_B(x) \cdot p_D(x)$.

Proof. Let s^A_i, s^B_i, s^D_i be the coefficients of the characteristic polynomials (as given by (1)) of A, B, D, respectively. We want to show by induction on i that

$$s^A_i = \sum_{j+k=i} s^B_j s^D_k,$$

from which the claim of the lemma follows. The Basis Case: $s^A_0 = s^B_0 = s^D_0 = 1$. For the Induction Step, by definition and by the induction hypothesis, we have that $(i+1) \cdot s^A_{i+1}$ equals

$$= \sum_{j=0}^{i} (-1)^j s^A_{i-j} \operatorname{tr}(A^{j+1}) = \sum_{j=0}^{i} (-1)^j \left[\sum_{p+q=i-j} s^B_p s^D_q \right] \operatorname{tr}(A^{j+1})$$

and by the form of A (i.e., (11)):

$$= \sum_{j=0}^{i} (-1)^j \left[\sum_{p+q=i-j} s^B_p s^D_q \right] (\operatorname{tr}(B^{j+1}) + \operatorname{tr}(D^{j+1}))$$

to see how this formula simplifies, we divide it into two parts:

$$= \sum_{j=0}^{i} (-1)^j \left[\sum_{p+q=i-j} s_p^B s_q^D \right] \mathrm{tr}(B^{j+1}) + \sum_{j=0}^{i} (-1)^j \left[\sum_{p+q=i-j} s_p^B s_q^D \right] \mathrm{tr}(D^{j+1}).$$

Consider first the left-hand side. When $q = 0$, p ranges over $\{i, i-1, \ldots, 0\}$, and $j+1$ ranges over $\{1, 2, \ldots, i+1\}$, and therefore, by definition, we obtain $(i+1) \cdot s_{i+1}^B$. Similarly, when $q = 1$, we obtain $i \cdot s_i^B$, and so on, until we obtain $1 \cdot s_1^B$. Hence we have:

$$= \sum_{j=0}^{i} ((i+1) - j) \cdot s_{(i+1)-j}^B s_j^D \quad + \quad \sum_{j=0}^{i} (-1)^j \left[\sum_{p+q=i-j} s_p^B s_q^D \right] \mathrm{tr}(D^{j+1}).$$

The same reasoning, but fixing p instead of q on the right-hand side, gives us:

$$= \sum_{j=0}^{i} ((i+1) - j) \cdot s_{(i+1)-j}^B s_j^D \quad + \quad \sum_{j=0}^{i} s_j^B ((i+1) - j) \cdot s_{(i+1)-j}^D$$

which gives us $(i+1) \cdot \sum_{j+k=i+1} s_j^B s_k^D$ and finishes the proof. □

References

[1] David S. Dummit and Richard M. Foote. *Abstract Algebra.* Prentice Hall, 1991.

[2] Paul R. Halmos. *Linear algebra problem book.* The mathematical association of America, 1995.

[3] W. Keith Nicholson. *Linear algebra with applications.* PWS Publishing Company, 1995.

[4] Michael Soltys. Feasible proofs of matrix properties with Csanky's algorithm. In *Computer Science Logic (CSL'05)*, pages 493–508, 2005.

[5] Michael Soltys and Stephen Cook. The complexity of derivations of matrix identities. *Annals of Pure and Applied Logic*, 130(1–3):277–323, December 2004.

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

3 Appendix

LA: three sorts: indices $i, j, k, \ldots$; field elements $a, b, c, \ldots$; matrices $A, B, C, \ldots$

Equality Axioms

E1 $\rightarrow x = x$
E2 $x = y \rightarrow y = x$
E3 $(x = y \wedge y = z) \rightarrow x = z$
E4 $x_1 = y_1, \ldots, x_n = y_n \rightarrow fx_1 \ldots x_n = fy_1 \ldots y_n$
E5 $i_1 = j_1, i_2 = j_2, i_1 \leq i_2 \rightarrow j_1 \leq j_2$

Index Axioms

I6 $\rightarrow i + 1 \neq 0$
I7 $\rightarrow i * (j + 1) = (i * j) + i$
I8 $i + 1 = j + 1 \rightarrow i = j$
I9 $\rightarrow i \leq i + j$
I10 $\rightarrow i + 0 = i$
I11 $\rightarrow i \leq j, j \leq i$
I12 $\rightarrow i + (j + 1) = (i + j) + 1$
I13 $i \leq j, j \leq i \rightarrow i = j$
I14 $\rightarrow i * 0 = 0$
I15 $i \leq j, i + k = j \rightarrow j - i = k$
$i \not\leq j \rightarrow j - i = 0$
I16 $j \neq 0 \rightarrow \mathtt{rem}(i, j) < j$
$j \neq 0 \rightarrow i = j * \mathtt{div}(i, j) + \mathtt{rem}(i, j)$
I17 $\alpha \rightarrow \mathrm{cond}(\alpha, i, j) = i \quad \neg\alpha \rightarrow \mathrm{cond}(\alpha, i, j) = j$

Fields Axioms

F18 $\rightarrow 0 \neq 1 \wedge a + 0 = a$
F19 $\rightarrow a + (-a) = 0$
F20 $\rightarrow 1 * a = a$
F21 $a \neq 0 \rightarrow a * (a^{-1}) = 1$
F22 $\rightarrow a + b = b + a$
F23 $\rightarrow a * b = b * a$
F24 $\rightarrow a + (b + c) = (a + b) + c$
F25 $\rightarrow a * (b * c) = (a * b) * c$
F26 $\rightarrow a * (b + c) = a * b + a * c$
F27 $\alpha \rightarrow \mathrm{cond}(\alpha, a, b) = a$
$\neg\alpha \rightarrow \mathrm{cond}(\alpha, a, b) = b$

Axioms for Matrices

M28 $(i = 0 \vee \mathtt{r}(A) < i \vee j = 0 \vee \mathtt{c}(A) < j) \rightarrow \mathtt{e}(A, i, j) = 0$
M29 $\rightarrow \mathtt{r}(\lambda ij\langle m, n, t\rangle) = m \quad \rightarrow \mathtt{c}(\lambda ij\langle m, n, t\rangle) = n$
$1 \leq i, i \leq m, 1 \leq j, j \leq n \rightarrow \mathtt{e}(\lambda ij\langle m, n, t\rangle, i, j) = t$
M30 $\mathtt{r}(A) = 1, \mathtt{c}(A) = 1 \rightarrow \Sigma(A) = \mathtt{e}(A, 1, 1)$
M31 $\mathtt{r}(A) = 1, 1 < \mathtt{c}(A) \rightarrow \Sigma(A) = \Sigma(\lambda ij\langle 1, \mathtt{c}(A) - 1, A_{ij}\rangle) + A_{1\mathtt{c}(A)}$
M32 $\mathtt{c}(A) = 1 \rightarrow \Sigma(A) = \Sigma(A^t)$
M33 $1 < \mathtt{r}(A), 1 < \mathtt{c}(A) \rightarrow \Sigma(A) = \mathtt{e}(A, 1, 1) + \Sigma(\mathtt{R}(A)) + \Sigma(\mathtt{S}(A)) + \Sigma(\mathtt{M}(A))$
M34 $\mathtt{r}(A) = 0 \vee \mathtt{c}(A) = 0 \rightarrow \Sigma A = 0$

Rules

The usual **LK** rules for logical consequence.

Equality: $$\frac{\Gamma \rightarrow \Delta, \mathtt{e}(T, i, j) = \mathtt{e}(U, i, j) \quad \Gamma \rightarrow \Delta, \mathtt{r}(T) = \mathtt{r}(U) \quad \Gamma \rightarrow \Delta, \mathtt{c}(T) = \mathtt{c}(U)}{\Gamma \rightarrow \Delta, T = U}$$

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