Matrix Identities and the Pigeonhole Principle

Michael Soltys
Department of Computing and Software
McMaster University

Alasdair Urquhart
Department of Philosophy
University of Toronto

May 30, 2003

Abstract

We show that short bounded-depth Frege proofs of matrix identities, such as $PQ = I \supset QP = I$ (over the field of two elements), imply short bounded-depth Frege proofs of the pigeonhole principle. Since the latter principle is known to require exponential-size bounded-depth Frege proofs, it follows that the propositional version of the matrix principle also requires bounded-depth Frege proofs of exponential size.

1 Introduction

Elementary principles of linear algebra can be formulated as tautologies in propositional logic. This paper is concerned with assessing the complexity of proofs required for these tautologies in various proof systems. We are particularly interested in the principle that a one-sided inverse of a square matrix is also a two-sided inverse, that is to say, $PQ = I \supset QP = I$. We refer to this implication as "the inversion principle," or IP; we use IP_n to refer to the inversion principle restricted to $n \times n$ matrices.

Stephen A. Cook has suggested this principle as a tautology that may be hard to prove in conventional systems of propositional logic. Specifically, he has asked whether this principle has polynomial-size proofs in Frege systems (conventional textbook-style proof systems that use schematic axioms and rules). If in fact the principle requires super-polynomial size proofs in Frege systems then this would provide a separation result between such systems and extended Frege systems, that add to Frege systems the possibility of abbreviating complex formulas by single variables. This is because short proofs of the inversion principle exist if we allow the possibility of abbreviative definitions, and formalize the Gaussian Elimination algorithm ([8]), or the Cayley-Hamilton theorem ([4] and [7]). On the other hand, without such definitions, the usual proof apparently leads to formulas of exponential size.

In the first part of the paper, we show that the principle IP_n requires proofs of exponential size in a Frege system where the formulas are restricted to those of a fixed depth. The method of proof is to show that when written in an appropriate way as a formula of fixed depth (allowing unbounded fan-in conjunctions and disjunctions), the assumption IP_n, over the field of two elements, leads to a short proof of a form of the pigeonhole principle. Since this latter tautology requires exponentially large proofs in a bounded-depth Frege system, so does the inversion principle.

The short proof of the pigeonhole principle obviously carries over to the Frege system where we enlarge our class of bounded-depth formulas to include unbounded fan-in $\oplus$ (XOR) gates. It is conjectured in this case also that the pigeonhole principle, appropriately formulated, requires exponentially large proofs. If this conjecture is correct, then this implies that the inversion principle is unprovable in a proof system for elementary linear algebra formulated by the first author in his doctoral dissertation [7]. This proof system, called LA, is a quantifier-free logical theory that is strong enough to prove all the ring properties of matrices, but weak enough so that all its theorems (over the field of two elements) translate into families of tautologies with short bounded depth Frege proofs with $\oplus$. A weaker (but still interesting) result than that IP_n separates Frege and Extended Frege would be whether IP_n is independent of the theory LA. We present briefly the theory LA in section 3.

It is also unknown whether IP_n has quasi-polynomial size Frege proofs. Since we can compute inverses in the class NC^2 using Berkowitz's algorithm, it would seem natural to be able to conclude that NC^2-Frege can prove IP_n. However, we do not know how to prove the correctness of Berkowitz's algorithm in NC^2-Frege, and hence we do not know how to prove IP_n in NC^2 (i.e., with quasi-polynomial size Frege proofs). We have feasible proofs of the correctness of Berkowitz's algorithm, but this only gives polysize Extended Frege proofs of IP_n (see [4]).

The excellent recent text of Clote and Kranakis [3] covers the basic background on proof systems assumed in the paper.

2 A lower bound for the inversion principle

In this section, we prove an exponential lower bound for bounded-depth Frege proofs of the inversion principle over the two-element field.

The idea for the proof is the following: suppose that we can prove IP_n in polysize bounded-depth Frege. Then we could also prove PHP_n in polysize bounded-depth Frege, where PHP_n is formulated as follows: an injective mapping from a set with n elements to a set with n elements is necessarily surjective. To see this, assume that we have an injective mapping $P : [n] \longrightarrow [n]$, and let P_n be the $n \times n$ matrix where the (i, j) entry is 1 if $P(i) = j$, and zero otherwise. Then, since P is injective, and hence each column of P_n has at most one 1, $P_n P_n^t = I_n$ (P_n^t is the transpose of P_n). Using IP_n we conclude that $P_n^t P_n = I_n$ as well, so each column of P_n has at least one 1, and so P is surjective. In the rest of this section, we show how to formalize this argument with short bounded depth Frege proofs.

There is an immediate difficulty in that the most obvious formalization of the inversion principle results in formulas of unbounded depth. Let P, Q be $n \times n$ square $\{0,1\}$ matrices; then the matrix equation $PQ = I$ can be expressed by the family of n^2 Boolean equivalences: $\sum_{k=1}^{n} P_{ik}Q_{kj} \equiv \delta_{ij}$, where the summation is over the two-element field, and δ_{ij} is 1 if $i = j$, and 0 otherwise. However, if we expand the sum $x_1 \oplus \cdots \oplus x_n$ in the obvious way, using the equivalence $(x \oplus y) \equiv \neg(x \vee y) \vee \neg(\neg x \vee \neg y)$, then the result is a formula of unbounded depth. We can surmount this difficulty by introducing auxiliary variables (also called "extension variables").

For each i, j, we introduce $n - 1$ new variables E_{ij}^k, F_{ij}^k, and the defining equivalences $E_{ij}^k \equiv (E_{ij}^{k-1} \oplus P_{ik}Q_{kj})$, $F_{ij}^k \equiv (F_{ij}^{k-1} \oplus Q_{ik}P_{kj})$, for $k > 1$, and $E_{ij}^1 \equiv P_{i1}Q_{1j}$, $F_{ij}^1 \equiv Q_{i1}P_{1j}$. The matrix equation $PQ = I$ can then be written as the conjunction of these defining equivalences, together with the n^2 Boolean equivalences $E_{ij}^n \equiv \delta_{ij}$. With this formalization, the inversion principle IP_n can be written as a formula of bounded depth of size $O(n^3)$.

To formulate the Pigeonhole principle we use propositional variables P_{ij} to represent the statement $P(i) = j$, where P is a mapping. The antecedent of PHP_n is the conjunction of the following three formulas:

$$\bigwedge_{1\le i\le n} P_{i1} \vee \cdots \vee P_{in}; \quad \bigwedge_{1\le i\le n, 1\le k<j\le n} \neg P_{ik} \vee \neg P_{ij}; \quad \bigwedge_{1\le i<j\le n, 1\le k\le n} \neg P_{ik} \vee \neg P_{jk}$$

where the first formula states that the domain of P is $\{1, \ldots, n\}$, the second states that P is a function, and the third states that P is injective. The consequent of PHP_n is the formula:

$$\bigwedge_{1\le i\le n} P_{1i} \vee P_{2i} \vee \cdots \vee P_{ni},$$

which states that P is surjective.

The lower bound of this section is based on a fundamental lower bound for bounded depth proofs of the pigeonhole principle (presented as theorem 2.1 below). It is the result of a series of papers [2, 5, 6] improving on the original seminal result of Ajtai [1]. A simplified exposition of the lower bound is given in [9]; see also Clote and Kranakis [3, Chapter 5].

Theorem 2.1 *Let $\mathcal{F}$ be a Frege system and $d > 4$. Then for sufficiently large n every depth d proof in $\mathcal{F}$ of PHP_n must have size at least 2^{n^δ}, for $\delta < (1/5)^d$.*

We now define a substitution instance IP_n^σ of IP_n; our main lemma shows that PHP_n can be deduced efficiently from this instance. The formula IP_n^σ is obtained from IP_n by the substitution: $Q_{ij} \mapsto P_{ji}$, $E_{ij}^k \mapsto (P_{i1}P_{j1} \vee \cdots \vee P_{ik}P_{jk})$, $F_{ij}^k \mapsto (P_{1i}P_{1j} \vee \cdots \vee P_{ki}P_{kj})$. If A is a formula, then we denote the result of applying this substitution to A by A^σ.

Lemma 2.2 *There is a bounded depth Frege proof of PHP_n from IP^σ_n of size $O(n^5)$.*

Proof. We will argue that if we have IP^σ_n, and the antecedent of PHP_n, then we can conclude the antecedent of IP^σ_n. Once we have the antecedent of IP^σ_n, we can obtain the consequent of IP^σ_n, and hence the consequent of PHP_n.

Assume IP^σ_n, together with the antecedent of PHP_n. Our first goal is to deduce the antecedent of IP^σ_n. It should be clear to the reader that all of the proofs sketched below can be carried out in a relatively small fixed depth, though we do not compute this depth explicitly.

The abbreviation $[E^k_{ij} \equiv (E^{k-1}_{ij} \oplus P_{ik}Q_{kj})]^\sigma$ is given by:

$$(P_{i1}P_{j1} \vee \cdots \vee P_{ik}P_{jk}) \equiv ((P_{i1}P_{j1} \vee \cdots \vee P_{i(k-1)}P_{j(k-1)}) \oplus P_{ik}P_{jk})$$

and it can be proven from $\neg(P_{i1}P_{j1} \vee \cdots \vee P_{i(k-1)}P_{j(k-1)}) \vee \neg(P_{ik}P_{jk})$, which the reader can check is easily derivable from the antecedent of PHP_n.

Since $E^n_{ij} \equiv (P_{i1}P_{j1} \vee \cdots \vee P_{in}P_{jn})$, if $i = j$, then $E^n_{ij} \equiv 1$, and if $i \neq j$, then $\neg(P_{i1}P_{j1} \vee \cdots \vee P_{in}P_{jn})$ is derivable from $\neg P_{ik} \vee \neg P_{jk}$, so $E^n_{ij} \equiv 0$. Since $(PQ = I)^\sigma$ is $(E^n_{ij})^\sigma$, we have just shown the antecedent of IP_n.

The last part of the derivation consists in deriving the consequent of PHP_n. To obtain this, note that we now have $(QP = I)^\sigma$, so we have that $(F^n_{ii} \equiv 1)^\sigma$, or equivalently, we have $(F^n_{ii})^\sigma$, which is just $(P_{1i}P_{1i} \vee \ldots \vee P_{ni}P_{ni})$ from which we can easily obtain $(P_{1i} \vee \ldots \vee P_{ni})$. We repeat this for every i, and obtain the consequent of PHP_n. □

We can now prove our main theorem by combining the previous lower bound for the pigeonhole principle with the preceding lemma.

Theorem 2.3 *Let $\mathcal{F}$ be a Frege system. Then there is a constant c so that for $d > c$ and sufficiently large n, every depth d proof in $\mathcal{F}$ of IP_n must have size $2^{\Omega(n^\delta)}$, for $\delta < (1/5)^{d+1}$.*

Proof. Choose the constant $c > 3$ so that it is at least as big as the depth of the derivation constructed in Lemma 2.2. Let $d > c$, and in addition, let D be a depth d derivation of IP_n, of size s. Then the derivation D^σ obtained by replacing all steps A by A^σ is a depth $d + 1$ derivation of IP^σ_n of size $O(ns)$. By Lemma 2.2, we can convert D^σ into a depth $d + 1$ derivation of PHP_n by appending a derivation of size $O(n^5)$. By Theorem 2.1, D^σ must have size at least 2^{n^δ}, for $\delta < (1/5)^{(d+1)}$. Hence, $sn \geq (2^{n^\delta} - O(n^5))$, and therefore $s \geq 2^{n^\delta - \log_2 n} - O(n^4)$. We can choose a new δ, still $< (1/5)^{d+1}$, so that for sufficiently large n, the terms "$\log_2 n$" and "$O(n^4)$" are eliminated. This gives us the result. □

We can modify our definition of depth to allow unbounded fan-in $\oplus$ (XOR) gates, in addition to unbounded AND and OR gates. With this modification, it is not known whether the the pigeonhole principle PHP_n still requires exponentially large proofs, but this is generally conjectured to be the case. If this conjecture is correct, then Lemma 2.2 applies to show exponential lower bounds for the inversion principle in this case as well.

3 The theory LA

LA is a quantifier-free, three-sorted logical theory, which can prove the ring properties of matrices, such as commutativity of matrix addition or associativity of matrix products. The matrix principle IP_n can be stated in the language of LA, but we conjecture that it is not a theorem of LA.

The theorems of LA (over the field of two elements) can be translated into families of propositional tautologies with polynomial size bounded-depth Frege proofs, with $\oplus$ gates of unbounded fan-in, i.e., in $\mathrm{AC}^0[2]$-Frege. Thus, to show that IP_n is independent of LA, it would be sufficient to show that $\mathrm{AC}^0[2]$-Frege does not prove IP_n. Thus, if $\mathrm{AC}^0[2]$ does not prove PHP_n, we could use Lemma 2.2 to conclude that $\mathrm{AC}^0[2]$-Frege does not prove IP_n, and hence that IP_n is independent of LA.

In this section we give a brief description of LA; for a full treatment of LA see [7] and [4]. We also present other universal matrix identities, similar to IP_n, which can be proven equivalent to IP_n in LA (and hence, their propositional counterparts can be shown equivalent to IP_n in $\mathrm{AC}^0[2]$). We also state the general simulation result for the theorems of LA; that is, we will spell out in detail the assertion that all the theorems of LA can be translated into families of tautologies with short $\mathrm{AC}^0[2]$-Frege proofs, over the field $\mathbb{F}_2$.

The three sorts of LA are *indices*, *field elements* (or just elements of a *commutative ring*), and *matrices*. We shall denote index variables by i, j, k, field variables by a, b, c, and matrix variables by P, Q, R. We shall denote the formulas of LA by α, β. LA has the usual arithmetic function symbols for indices: addition, multiplication, subtraction, and also function symbols for division and remainder. There is also addition and multiplication for field elements, as well as additive and multiplicative inverses for field elements. When considering a commutative ring rather than a field, the multiplicative inverse is not added. Thus, if m, n are index terms, then so are $(m{+_i}n), (m{*_i}n), (m{-_i}n), div(m,n), rem(m,n)$ (where the subscript "$_i$" indicates that these are index operations), and if t, u are terms of type field, then so are $(t +_f u), (t *_i u), (-_i t), (t^{-1})$.

If T is a term of type matrix, then $r(T), c(T)$ are terms of type index which denote the number of rows and columns of T, respectively, and $\Sigma(T)$ is a term of type field that denotes the sum of all the entries of T, and if m, n are terms of type index, then $e(m, n, T)$ is a term of type field which denotes the (m, n) entry of the matrix T. All matrix variables $A, B, C, P, Q, \ldots$ are matrix terms. We construct new matrices using some rudimentary λ-calculus: if m, n are terms of type index, and t is a term of type field, then $\lambda ij\langle m, n, t\rangle$ is a *constructed term of type matrix* (note that the index variables i, j cannot occur free in m, n). That is, $r(\lambda ij\langle m, n, t\rangle) = m$, $c(\lambda ij\langle m, n, t\rangle) = n$, and $e(i, j, \lambda ij\langle m, n, t\rangle) = t$.

If m, n, t, u, T, U are terms, then $(m \leq n), (m = n), (t = u), (T = U)$ are atomic formulas of the appropriate kind (index, index, field, matrix, respectively). We build general formulas in the usual way: if α, β are formulas, then so are: $(\neg\alpha)$, $(\alpha \vee \beta)$ and $(\alpha \wedge \beta)$.

Finally, if α is a formula where all the atomic subformulas are of type index, then $\mathrm{cond}_i(\alpha, m, n)$ and $\mathrm{cond}_f(\alpha, t, u)$ are terms of type index and field, respectively, and the idea is that $\mathrm{cond}_i(\alpha, m, n)$ is m if α is true, and n otherwise, and similarly for cond_f. The restriction that all the atomic subformulas of α are of type index is there because in the translations into propositional formulas, all the free index variables get values, and therefore, α will become true or false.

All the usual axioms for equality are in LA. We have the usual axioms of Robinson's arithmetic Q in LA together with axioms defining div, rem, and cond, for elements of type index. The axioms for field elements are the usual field axioms. The axioms for matrices define the behavior of constructed matrices, and define the function Σ recursively (first on row matrices, and then on general matrices). As LA is a quantifier-free theory, these are really *axiom-schemes*, since we allow any substitution of terms for variables.

To prove theorems in LA, we have the usual Frege rules for propositional consequence, and a rule for induction on indices, and a rule for concluding equality of matrices. The *induction rule* is: $\alpha(i) \supset \alpha(i+1) \vdash \alpha(0) \supset \alpha(n)$, note that i must be an index variable, and must not occur free on the right-hand side of the rule.

It turns out that LA is "strong enough" to prove all the ring properties of the set of matrices (i.e., properties such as the associativity of matrix multiplication, commutativity of matrix addition, etc.). In other words, all the ring properties of matrices, expressed in the language of LA, *are* in LA. For the formal derivations that place these ring properties in LA see [7, §3.1].

On the other hand, LA is "weak enough," so that all the theorems of LA have $\mathrm{AC}^0[2]$-Frege proofs; we state this more precisely below with two theorems, but essentially this means that all the ring properties of matrices can be proven with $\mathrm{AC}^0[2]$-Frege. We recall that LA is field (commutative ring, if the multiplicative inverse function is ignored) independent. Thus, when translating the theorems of LA into a propositional proof system, we have to fix the field. In the context of this paper, we only consider the field of two elements $\{0, 1\}$. If the underlying field were $\mathbb{F}_p$, then the theorems of LA would translate into $\mathrm{AC}^0[p]$-Frege proofs.

The general method for translations is given in [7, §7]. There, we give a natural recursive procedure that takes as input a formula α over the language of LA, and produces a family of tautologies $\{\|\alpha\|_\sigma\}$ parametrized by σ. The important properties of the translation are given in the following theorem. Let $|\sigma|$ be the largest value in the assignment σ. Let $\|\alpha\|_\sigma$ be the translation of α into a family of propositional tautologies, parametrized by σ; for each σ we get a tautology of different size. (See [7, §7] for the proof.)

Theorem 3.1 *If α is a formula over the language of LA, then, there exists a polynomial p_α and a constant d_α such that for every σ, the size of $\|\alpha\|_\sigma$ is bounded by $p_\alpha(|\sigma|)$, and the depth of $\|\alpha\|_\sigma$ is bounded by d_α. Furthermore, if α is a true formula (in the standard model) then, the propositional formula $\|\alpha\|_\sigma$ is a tautology. Furthermore, if α is a theorem of LA, then, there exists a polynomial q_α and a positive integer d_α such that for every σ, $\|\alpha\|_\sigma$ has an $AC^0[2]$-Frege derivation $\pi_{\alpha,\sigma}$ such that the size of $\pi_{\alpha,\sigma}$ is bounded by $q_\alpha(|\sigma|)$ and the depth of $\pi_{\alpha,\sigma}$ is bounded by the constant d_α.*

Thus, it is possible to state matrix principles, such as IP_n in the language of LA, but we conjecture that LA is too weak to prove them. However, LA is strong enough to prove the equivalence of a host of "hard" matrix identities, such as: $(PQ = I \wedge PR = I) \supset Q = R$, $PQ = I \supset (PR \neq 0 \vee R = 0)$, $PQ = I \supset P^tQ^t = I$. Since all the theorems of LA translate into families of propositional tautologies with short, bounded-depth, Frege proofs (with $\oplus$), i.e., $\mathrm{AC}^0[2]$-Frege proofs, if we could show that IP_n, or *any* of the above matrix identities, is not provable in $\mathrm{AC}^0[2]$-Frege, we would have shown that they are *all* independent of LA.

Therefore, showing that PHP_n is not provable in $\mathrm{AC}^0[2]$ would allow us to conclude that IP_n, and the above three matrix identities, as well as any universal matrix identity equivalent in LA to IP_n, are all independent of LA. In other words, we would have shown that LA can prove ring properties of matrices, but it is too weak to prove matrix identities which express properties of inverses.

References

[1] Miklós Ajtai. The complexity of the pigeonhole principle. In *Proceedings of the 29th Annual IEEE Symposium on the Foundations of Computer Science*, pages 346–355, 1988.

[2] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, Pavel Pudlák, and Alan Woods. Exponential lower bounds for the pigeonhole principle. In *Proceedings of the 24th Annual ACM Symposium on theory of computing*, pages 200–220, 1992.

[3] Peter Clote and Evangelos Kranakis. *Boolean Functions and Computation Models.* Springer-Verlag, 2002.

[4] Stephen A. Cook and Michael Soltys. The proof complexity of linear algebra. In *Seventeenth Annual IEEE Symposium on Logic in Computer Science (LICS 2002)*, 2002.

[5] Jan Krajíček, Pavel Pudlák, and Alan Woods. Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle. *Random Structures and Algorithms*, 7:15–39, 1995.

[6] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Computational Complexity*, 3:97–140, 1993.

[7] Michael Soltys. *The Complexity of Derivations of Matrix Identities.* PhD thesis, University of Toronto, 2001.

[8] Michael Soltys. Extended Frege and Gaussian elimination. *Bulletin of the Section of Logic, Polish Academy of Sciences*, 31(4):189–206, 2002.

[9] Alasdair Urquhart and Xudong Fu. Simplified lower bounds for propositional proofs. *Notre Dame Journal of Formal Logic*, 37:523–544, 1996.