Permanent and Determinant*

Joachim von zur Gathen
Department of Computer Science
University of Toronto
Toronto, Ontario M5S 1A4, Canada

Submitted by Hans Schneider

ABSTRACT

The $n \times n$ permanent is not a projection of the $m \times m$ determinant if $m \leqslant \sqrt{2}\, n - 6\sqrt{n}$.

1. INTRODUCTION

The definitions of permanent and determinant look very similar:

$$\operatorname{per} x_{ij} = \sum_{\sigma \in \mathrm{Sym}_n} x_{1\sigma_1} \cdots x_{n\sigma_n}$$

differs from $\det x_{ij}$ only in the signs of the summands. However, while Gaussian elimination provides an efficient way of calculating the determinant, no fast algorithm is known for the permanent. (We assume an arbitrary ground field of characteristic different from two, since otherwise $\operatorname{per} x_{ij} = \det x_{ij}$.) Evidence for the difficulty of computing the permanent was given in Valiant's (1979a) theory of p-completeness, an arithmetic analogue of the Boolean theory of P versus NP [see von zur Gathen (1987) for an exposition].

*Part of this work was done while the author was visiting Universität Zürich, and supported by Schweizerischer Nationalfonds, grant 2175-0.83, and by NSERC, grant 3-650-126-40. A preliminary version appears in *Proceedings of the 27th Annual IEEE Symposium on Foundations of Computer Science*, Toronto, Ontario, 1986, pp. 398–401.

The determinant can be computed in polynomial time, but the permanent is "p-complete": a polynomial-time algorithm for the permanent would imply one for a host of problems which have so far withstood attempts to find fast algorithms. Valiant's hypothesis is the conjecture that no such fast algorithms exist, in particular no arithmetic algorithms for the $n \times n$ permanent using constants from the ground field, indeterminates, and $n^{O(1)}$ arithmetic operations $+$, $-$, $*$, $/$. One of the motivations for Valiant's theory is the hope that the powerful tools of algebra may allow us to solve problems which are very hard in the Boolean context, maybe even Valiant's arithmetic analogue of Cook's hypothesis P $\neq$ NP.

The Boolean problem of computing the permanent of a matrix with integer entries is #P-complete (Valiant 1979b).

The $n \times n$ permanent is said to be a projection of the $m \times m$ determinant if there exists an $m \times m$ matrix f whose entries are constants and indeterminates x_{ij} $(1 \leqslant i, j \leqslant n)$ such that $\operatorname{per} x_{ij} = \det f$. Let us denote by $p(n)$ the smallest such m. Valiant proves $p(n) = O(n^2 2^n)$. Valiant's hypothesis would follow from

$$p(n) = 2^{(\log n)^{\omega(1)}}.$$

$p(n) > n$ is trivial. The main result of the present paper is the first nontrivial lower bound for this problem, showing that $p(n) \geqslant \sqrt{2}\, n - 6\sqrt{n}$ over an infinite field of characteristic different from two. The author had obtained $p(n) > 1.06n - 1$; the stated bound is due to Babai and Seress (1987).

The question of what kind of relations exist between permanent and determinant, in particular whether the permanent can be expressed as the determinant of a matrix, is a classical mathematical problem. Szegő (1913), answering a question posed by Pólya (1913), showed that for $n \geqslant 3$, there is no way of generalizing

$$\operatorname{per}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \det\begin{bmatrix} x_{11} & -x_{12} \\ x_{21} & x_{22} \end{bmatrix},$$

i.e., of affixing $\pm$ signs to the indeterminate entries x_{ij} such that

$$\operatorname{per} x_{ij} = \det(\pm x_{ij}).$$

In view of this question, we consider "$\pm$-projections" f with $\operatorname{per} x = \det f$ as above, but where now constants, indeterminates x_{ij}, and also $-x_{ij}$ are allowed. $p_{\pm}(n)$ is the minimal m with this property. Clearly $p_{\pm}(n) \leqslant p(n)$.

Marcus and Minc (1961) proved that one cannot relate certain permanental and determinantal functions by linear mappings. In particular, for $n \geqslant 3$,

there are no linear forms f_{kl} in indeterminates x_{ij} $(1 \leqslant i, j, k, l \leqslant n)$ such that $\operatorname{per} x_{ij} = \det f_{kl}$. The methods of this paper yield an easy proof of this result, generalized to arbitrary infinite fields of characteristic different from two, and also allowing affine linear forms with nonzero constant terms.

A general background on permanents is given in Minc (1978).

The paper is organized as follows. In Section 2, we give bounds on the dimension (in the sense of algebraic geometry) of the singular locus of the permanent and determinant polynomials. In Section 3, we derive a criterion on mappings that transform the permanent into the determinant. The result of Marcus and Minc follows immediately. As an aside, absolute irreducibility of the permanent is a corollary. Applying the criterion, a combinatorial argument proves in Section 4 that $p_{\pm}(n) > \sqrt{2}\, n - 6\sqrt{n}$.

We note that the combinatorial argument can be applied directly to prove lower bounds for $p(n)$, without using the geometry of Sections 2 and 3 (see the preliminary version). However, that approach seems a dead end, while it remains open whether the present method can yield better lower bounds.

2. THE SINGULAR LOCUS OF DETERMINANT AND PERMANENT

Throughout the paper, F is a field of characteristic different from two, $n \in \mathbb{N}$, and x_{ij} are indeterminates over F for $1 \leqslant i, j \leqslant n$. We use elementary notions from algebraic geometry, as e.g. in Shafarevich (1974, Chapter I). For simplicity, we assume F algebraically closed in this section. We say that the $n \times n$ matrix $x = (x_{ij})_{1 \leqslant i, j \leqslant n}$ consists of the coordinates on the ring $F^{n \times n}$ of $n \times n$ matrices over F. We also let $\mathbf{x} = \{x_{11}, x_{12}, \ldots, x_{nn}\}$, so that $\operatorname{per} x, \det x \in F[\mathbf{x}]$. If $f \in F[y_1, \ldots, y_n]$ is square-free, then the singular locus $\operatorname{sing} Y$ of the hypersurface

$$Y = \{f = 0\} = \{a \in F^n : f(a) = 0\} \subseteq F^n$$

is the closed subvariety of F^n defined as

$$\operatorname{sing} Y = \left\{ a \in Y : \frac{\partial f}{\partial y_1}(a) = \cdots = \frac{\partial f}{\partial y_n}(a) = 0 \right\}.$$

Let $I, J \subseteq \{1, \ldots, n\}$, R a ring, and $u \in R^{n \times n}$ be an $n \times n$ matrix over R. We denote by $u(I|J)$ the $(n - \#I) \times (n - \#J)$ matrix obtained from u by deleting the rows from I and the columns from J. We also write $u(i|J)$ and

$u(i, j|J)$ if $I = \{i\}$ and $I = \{i, j\}$, respectively; similarly for J. We let

$$D_n = \{\det x = 0\} \subseteq F^{n\times n},$$

$$P_n = \{\operatorname{per} x = 0\} \subseteq F^{n\times n}.$$

Since $\partial \det x/\partial x_{ij} = (-1)^{i+j}\det(x(i|j))$, and similarly for per, we have

$$\begin{aligned}\operatorname{sing} D_n &= \{a \in F^{n\times n} : \forall i, j \leqslant n \ \det a(i|j) = 0\}\\ &= \{a \in F^{n\times n} : \operatorname{rank} a \leqslant n-2\},\\ \operatorname{sing} P_n &= \{a \in F^{n\times n} : \forall i, j \leqslant n \ \operatorname{per} a(i|j) = 0\}.\end{aligned}$$

LEMMA 2.1. *Let F be algebraically closed, $n \geqslant 2$. Then* $\operatorname{sing}(D_n)$ *is irreducible of dimension* $n^2 - 4$.

Proof. For $1 \leqslant i < j \leqslant n$, let

$$S_{ij} = \{a \in F^{n\times n} : \text{rows } i \text{ and } j \text{ of } a \text{ are linearly dependent on the other rows of } a\}.$$

Then $S_{n-1,n}$, e.g., is the image of the mapping

$$\phi : F^{(n-2)\times n} \times F^{n-2} \times F^{n-2} \to F^{n\times n},$$

$$(b, c, d) \mapsto \begin{bmatrix} b \\ \sum c_k b_{k*} \\ \sum d_k b_{k*} \end{bmatrix},$$

where b_{k*} is the kth row of b. The generic fibre of ϕ consists of one point, and therefore each S_{ij} is irreducible of dimension $n^2 - 4$. Furthermore,

$$\operatorname{sing} D_n = \bigcup_{1\leqslant i<j\leqslant n} S_{ij},$$

and for any $i < j$, $S_{ij} \cap S_{n,n-1}$ is a dense open subset of $S_{n,n-1}$. It follows that $\operatorname{sing} D_n$ equals the Zariski closure of $S_{n-1,n}$. ∎

This lemma is also valid in characteristic two.

EXAMPLE 2.2. In this example we determine $\operatorname{sing} P_3$. First note that for any n, the group G_n consisting of row permutations, column permutations and transposition on $F^{n\times n}$ leaves $\operatorname{sing} P_n$ invariant. Let

$$U = \left\langle \begin{bmatrix} * & 0 & 0 \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix} \right\rangle \subseteq F^{3\times 3},$$

$$V = \left\langle \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 0 \end{bmatrix} : \operatorname{per}\begin{bmatrix} * & * \\ * & * \end{bmatrix} = 0 \right\rangle \subseteq F^{3\times 3},$$

$$W = \bigcup_{\sigma\in G_3} [\sigma(U) \cup \sigma(V)].$$

Here, $*$ denotes an arbitrary entry from F. W consists of $6+9=15$ irreducible components of dimension 3. Clearly $W \subseteq \operatorname{sing} P_3$, and we prove that equality holds. Let $a \in \operatorname{sing} P_3$. We can assume that some entry of a is nonzero, and hence that $a_{11} \neq 0$. Then

$$a_{22} = \frac{-a_{12}a_{21}}{a_{11}},$$

and similarly a_{ij} for $i, j \in \{2,3\}$ are rational functions of a_{i1} and a_{1j}, $1 \leqslant i, j \leqslant 3$.

After substitution and multiplication by $-a_{11}/2$, the remaining five equations give

$$\begin{aligned}0 &= a_{12}a_{13}a_{21} = a_{12}a_{13}a_{31} = a_{12}a_{21}a_{31}\\ &= a_{13}a_{21}a_{31} = a_{12}a_{13}a_{21}a_{31}/a_{11}.\end{aligned}$$

All the solutions are contained in W. ∎

LEMMA 2.3. *Let $n \geqslant 3$, and F be algebraically closed. Then every irreducible component of* $\operatorname{sing} P_n$ *has dimension at most* $n^2 - 5$.

Proof. Let $C \subseteq F^{n\times n}$ be an irreducible component of $\operatorname{sing} P_n$. If all $(n-2)\times(n-2)$ permanents vanish on C, the claim follows by induction.

After possibly reordering rows and columns, we can assume that with $J=\{n-1,n\}$ and $g=\operatorname{per}x(J|J)$ we have $g\restriction C\neq 0$. For $1\leqslant i,j\leqslant n$, $f_{ij}=\operatorname{per}x(i|j)$ vanishes on C. Then e.g. developing f_{nn} along the $(n-1)$st column gives

$$f_{nn}=\sum_{1\leqslant i\leqslant n-2} x_{i,n-1}\operatorname{per}x(i,n|J)+x_{n-1,n-1}g.$$

Thus $x_{n-1,n-1}$ is a rational function on C of the n^2-4 variables x_{ij} with $i\notin J$ or $j\notin J$. Similarly, we can use $f_{n-1,n-1}$, $f_{n-1,n}$, and $f_{n,n-1}$ to express x_{nn}, $x_{n,n-1}$, and $x_{n-1,n}$, respectively, as rational functions on C of the remaining n^2-4 variables. This proves $\dim C\leqslant n^2-4$, and it remains to find a nonzero polynomial h in these n^2-4 variables that vanishes on C. Consider

$$\begin{aligned}h&=gf_{n-2,n}-f_{n-1,n}\operatorname{per}x(n-2,n|J)\\&\quad-f_{nn}\operatorname{per}x(n-2,n-1|J)\\&=\sum_{1\leqslant i\leqslant n-3} x_{i,n-1}[\operatorname{per}x(i,n-2|J)\operatorname{per}x(n-1,n|J)\\&\qquad-\operatorname{per}x(i,n-1|J)\operatorname{per}x(n-2,n|J)\\&\qquad-\operatorname{per}x(i,n|J)\operatorname{per}x(n-2,n-1|J)]\\&\quad-2x_{n-2,n-1}\operatorname{per}x(n-2,n-1|J)\operatorname{per}x(n-2,n|J).\end{aligned}$$

The coefficient of $x_{n-2,n-1}$ in h is

$$-2\operatorname{per}x(n-2,n-1|J)\operatorname{per}x(n-2,n|J)\neq 0,$$

and therefore h is not the zero polynomial. Since $h\restriction C=0$, the claim follows. ∎

REMARK 2.4. We have not determined sing P_n, or its dimension. One can show that e.g. the matrices with last two columns equal to zero form a component of sing P_n, of dimension n^2-2n. The set of 4×4 matrices of the form

$$\begin{bmatrix}*&*&0&0\\*&*&0&0\\0&0&*&*\\0&0&*&*\end{bmatrix},$$

with the two 2×2 permanents equal to zero, forms a component C of sing P_4, of dimension 6. Each component of sing P_4 outside the orbit of C under G_4 has dimension 8, and contains at least one zero row or column. To generalize C, one takes two diagonal squares, say of sizes $s\times s$ and $(n-s)\times(n-s)$, and sets the corresponding two permanents equal to zero and all entries outside the squares equal to zero. The resulting dimension is $s^2+(n-s)^2-2$.

3. POLYNOMIAL MAPPINGS

The following theorem is our tool for proving lower bounds on functions relating the permanent and determinant.

THEOREM 3.1. *Let F be an infinite field of characteristic different from two, $m,n\in\mathbb{N}$, $n\geqslant 3$, x coordinates on $F^{n\times n}$, and*

$$f\colon F^{n\times n}\to F^{m\times m}$$

a polynomial mapping such that $\operatorname{im}f\cap\operatorname{sing}D_m\neq\emptyset$. *Then* $\operatorname{per}x\neq\det f$.

Proof. We first assume that F is algebraically closed, $\operatorname{per}x=\det f$, and let $y=(y_{kl})_{1\leqslant k,l\leqslant m}$ be coordinates on $F^{m\times m}$. Thus $D_m=\{\det y=0\}$. By assumption, f is given by polynomials:

$$f=(f_{kl})_{1\leqslant k,l\leqslant m}\in(F[\mathbf{x}])^{m\times m}.$$

(For the application in Section 4, it is sufficient to consider affine linear f.) For $a\in F^{n\times n}$ and $1\leqslant i,j\leqslant n$, we have

$$\begin{aligned}\operatorname{per}a(i|j)&=\left(\frac{\partial\operatorname{per}x}{\partial x_{ij}}\right)(a)=\left(\frac{\partial\det f}{\partial x_{ij}}\right)(a)\\&=\left(\sum_{1\leqslant k,l\leqslant m}\frac{\partial\det y}{\partial y_{kl}}(f)\frac{\partial f_{kl}}{\partial x_{ij}}\right)(a)\\&=\sum_{1\leqslant k,l\leqslant m}(-1)^{k+l}\det f(a)(k|l)\cdot\frac{\partial f_{kl}}{\partial x_{ij}}(a).\end{aligned}$$

If $f(a) \in \operatorname{sing} D_m$, then each summand vanishes, and therefore $\operatorname{per} a(i|j) = 0$ and $a \in \operatorname{sing} P_n$. Letting $S = f^{-1}(\operatorname{sing} D_m)$, we have shown that $S \subseteq \operatorname{sing} P_n$. If $S \neq \varnothing$, then $\dim S \geqslant n^2 - m^2 + (m^2 - 4) = n^2 - 4$ by Lemma 2.1 and the theorem on the dimension of fibres (Shafarevich 1974, Chapter I, Section 3). Therefore Lemma 2.3 implies that $S = \varnothing$; the claim is proven for algebraically closed fields.

If F is an arbitrary infinite field and K an algebraic closure of F, then f defines a polynomial mapping $\bar{f}: K^{n\times n} \to K^{m\times m}$. The theorem for $\bar{f}$ implies that for f. ∎

The theorem as stated also holds over a finite field F, since $\operatorname{per} x = \det f$ would be valid over arbitrary extension fields of F. However, it is more relevant to consider the condition

$$\forall a \in F^{n\times n} \qquad \operatorname{per} a = \det f(a),$$

and this may not extend to larger fields.

REMARK 3.2. The theorem is a special case of the following situation. Let $f: X \to Y$ be a morphism of smooth varieties, g a regular function on Y with $g \circ f \neq 0$, $V = \{g = 0\} \subseteq Y$, $W = \{g \circ f = 0\} = f^{-1}V \subseteq X$, and

$$T = \left\{ a \in W : \forall i \ \frac{\partial g \circ f}{\partial t_i}(a) = 0 \right\},$$

where the t_i form a system of local parameters on X at a. Then $f^{-1}\operatorname{sing} V \subseteq T$. Note that $\operatorname{sing} W \subseteq T$ may be a proper inclusion, if $g \circ f$ is not square-free: for $f: F \to F^2$ with $f(a) = (a, a)$ and $g = xy$, where x, y are the coordinates on F^2, we have $W = T = \{0\}$ and $\operatorname{sing} W = \varnothing$.

Marcus and Minc (1961) deal with representations of the permanent as linear combinations of determinants of linear forms. As a corollary, they obtain that for a field F of characteristic zero and $n \geqslant 3$, the $n \times n$ permanent is not the $n \times n$ determinant of some linear forms, i.e.,

$$\forall n \geqslant 3 \quad \forall f \in \left(\sum_{i,j} F x_{ij} \right)^{n\times n} \qquad \operatorname{per} x \neq \det f.$$

This is a trivial consequence of Theorem 3.1, since for any such $f: F^{n\times n} \to$

$F^{n\times n}$ we have $f(0) = 0 \in \operatorname{sing} D_m$. We generalize this statement to arbitrary infinite fields, and also allow constants in the linear forms, in order to include the "$\pm$-projections" considered in Section 4.

THEOREM 3.3. *Let F be an infinite field of characteristic different from two, $n \geqslant 3$, x coordinates on $F^{n\times n}$, and $f: F^{n\times n} \to F^{n\times n}$ affine linear. Then* $\operatorname{per} x \neq \det f$.

Proof. We split $f = g + h$ into its constant and linear parts, with $g \in F^{n\times n}$ and

$$h \in \left(\sum_{1 \leqslant i,j \leqslant n} F x_{ij} \right)^{n\times n},$$

and distinguish two cases.

Case 1: $\exists a \in F^{n\times n} \setminus \{0\}$ $h(a) = 0$. We choose $a \in F^{n\times n}$ with $h(a) = 0$ and $a_{11} \neq 0$, after possibly permuting rows and columns of x. Let $b \in F^{n\times n}$ be such that

$$b = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ & & * & \end{bmatrix},$$

$$a + b = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & 1 & & 0 \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix},$$

i.e., $b_{1i} = 0$ for $1 \leqslant i \leqslant n$, and $(a + b)_{ij} = \delta_{ij}$ for $2 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$. Then either $\operatorname{per} b \neq \det f(b)$, or else

$$\begin{aligned} \operatorname{per}(a + b) &= a_{11} \\ &\neq 0 = \operatorname{per} b = \det[g + h(b)] \\ &= \det[g + h(a) + h(b)] = \det[g + h(a + b)] = \det f(a + b). \end{aligned}$$

Case 2. $\forall a \in F^{n\times n} \setminus \{0\}$ $h(a) \neq 0$. Then f is an affine linear automorphism of $F^{n\times n}$, and $\operatorname{im} f = F^{n\times n}$. In particular, $0 \in \operatorname{im} f$, and the claim follows by Theorem 3.1. ∎

We note as an easy consequence the well-known absolute irreducibility of the permanent; this will not be used in the sequel.

THEOREM 3.4. *Let F be any field, $n \geq 1$, and x coordinates on $F^{n\times n}$. Then $\operatorname{per} x \in F[\mathbf{x}]$ is absolutely irreducible.*

Proof. We can assume that F is algebraically closed, since absolute irreducibility means irreducibility over an algebraic closure. If $\operatorname{char} F = 2$, then $\operatorname{per} x = \det x$ is irreducible [see e.g. van der Waerden (1970); that approach can also be used to prove Theorem 3.4]. For characteristic different from two, we can assume $n \geq 3$. Let $g, h \in F[\mathbf{x}]$ with $\operatorname{per} x = gh$. Consider the polynomial mapping $f: F^{n\times n} \to F^{2\times 2}$ given by

$$f = \begin{bmatrix} g & 0 \\ 0 & h \end{bmatrix} \in F[\mathbf{x}]^{2\times 2}.$$

Then $\operatorname{per} x = \det f$. Since $\operatorname{per} x$ is homogeneous, also g and h are homogeneous. If both g and h are nonconstant, then $g(0) = h(0) = 0$, and $f(0) = 0$, contradicting Theorem 3.1. ∎

4. THE PERMANENT AS PROJECTION OF THE DETERMINANT

We first recall Valiant's (1979a) notion of projection in our case.

DEFINITION 4.1. *Let F be a field, $n \in \mathbb{N}$, x coordinates on $F^{n\times n}$, $\mathbf{x} = \{x_{11}, x_{12}, \ldots, x_{nn}\}$, and*

$$p(n) = \min\{m \in \mathbb{N} : \exists f \in (F \cup \mathbf{x})^{m\times m} \ \operatorname{per} x = \det f\}.$$

If f is as above, then we say that the $n \times n$ permanent is a projection of the $m \times m$ determinant. Let $\pm\mathbf{x} = \mathbf{x} \cup \{-x_{11}, -x_{12}, \ldots, -x_{nn}\}$, and

$$p_{\pm}(n) = \min\{m \in \mathbb{N} : \exists f \in (F \cup \pm\mathbf{x})^{m\times m} \ \operatorname{per} x = \det f\}.$$

We call such an f a $\pm$-projection.

Clearly $p_{\pm}(n) \leq p(n)$, and $p_{\pm}(2) = 2 < p(2)$ (if $\operatorname{char} F \neq 2$). Valiant proves $p(n) = O(n^2 2^n)$. The interest in lower bounds on $p(n)$ stems from

the fact that

$$p(n) = 2^{(\log n)^{\omega(1)}}$$

implies Valiant's hypothesis, the arithmetic analogue of Cook's hypothesis P ≠ NP. Szegő [1913] showed $p_{\pm}(n) > n$; Theorem 4.4 improves this to $p_{\pm}(n) > \sqrt{2}\,n - 6\sqrt{n}$.

If we define p_{lin} and p_{aff} similarly, by allowing f to consist of linear or affine linear polynomials respectively, then Marcus and Minc (1961) prove $p_{\text{lin}}(n) > n$, and Theorem 3.3 reads $p_{\text{lin}} \geq p_{\text{aff}} > n$.

For the remainder of this section we fix the following notation. F is an infinite field of characteristic different from two, $m \geq n \geq 3$, x consists of coordinates on $F^{n\times n}$, $f \in (F \cup \pm\mathbf{x})^{m\times m}$, and $\operatorname{per} x = \det f$.

DEFINITION 4.2. A free square for f consists of four indices $1 \leq k_1 < k_2 \leq m$, $1 \leq l_1 < l_2 \leq m$ such that $f_{k_r l_s} \in \pm\mathbf{x}$, say

$$f_{k_r l_s} = \pm x_{i_{rs} j_{rs}} \in \pm\mathbf{x}$$

for $r, s \in \{1,2\}$; each such $f_{k_r l_s}$ occurs only once in f; $-f_{k_r l_s}$ does not occur in f; and $i_{11} \neq i_{22}$, $j_{11} \neq j_{22}$.

LEMMA 4.3. *f has no free square.*

Proof. We assume that k_1, k_2, l_1, l_2 form a free square for f. After possibly permuting rows and columns in f (so that $\operatorname{per} x = \pm\det f$), we can assume that $k_1 = l_1 = 1$, $k_2 = l_2 = 2$, so that

$$f = \left[\begin{array}{cc|c} f_{11} & f_{12} & g_1 \\ f_{21} & f_{22} & \\ \hline g_2 & & g_3 \end{array}\right],$$

where each f_{kl} is some $\pm x_{i_{kl} j_{kl}}$ for $k, l \in \{1,2\}$, and these variables do not occur in g_1, g_2, g_3. As a further simplification, we may also assume $f_{ij} = \pm x_{ij}$, say $f_{ij} = \epsilon_{ij} x_{ij}$ with $\epsilon_{ij} \in \{-1, 1\}$ for $i, j \in \{1,2\}$. We show that we can make $d = \det g_3$ nonzero by appropriate substitutions for x_{ij}, and then adjust the four variables in the top left corner so that the resulting matrix has rank $m - 2$. Let $S = \{x_{11}, x_{12}, x_{21}, x_{22}\}$ consist of the four special variables, and $z = f_{11} f_{22}$. Thus all entries of g_1, g_2, g_3 are from $T = (F \cup \pm\mathbf{x}) \setminus \pm S$. The coefficient of z in $\det f$ is $\pm d$, and the coefficient of z in $\operatorname{per} x$ is nonzero, using the last condition in Definition 4.2. Thus $d \neq 0$. The two rows

of $g_1 \in T^{2\times(m-2)}$ are linear combinations of the rows of $g_3 \in T^{(m-2)\times(m-2)}$, say

$$f_{ij} = \sum_{3 \leqslant k \leqslant n} \frac{u_{ik}}{d} f_{kj}$$

for $1 \leqslant i \leqslant 2 < j \leqslant n$, with $u_{ik} \in F[\mathbf{x} \setminus S]$.

We choose values a' in F for the $n^2 - 4$ indeterminates in $\mathbf{x} \setminus S$ such that $d(a') \neq 0$, and set

$$a_{ij} = \epsilon_{ij} \sum_{3 \leqslant k \leqslant n} \left(\frac{u_{ik}}{d} f_{kj} \right)(a') \in F$$

for $i, j \in \{1,2\}$. This completes a' to $a \in F^{n\times n}$ with rank $f(a) = m - 2$, contradicting Theorem 3.1. ∎

The author had originally used this lemma to show $p_{\pm}(n) > 1.06n - 1$. The following improvement is due to Babai and Seress (1987).

THEOREM 4.4. *Let F be an infinite field of characteristic different from two, and $n \in \mathbb{N}$. Then $p_{\pm}(n) > \sqrt{2}\, n - 6\sqrt{n}$.*

Proof. Suppose that $f \in (F \cup \pm\mathbf{x})^{m\times m}$ is a matrix with per $x = \det f$. We show that $m > \sqrt{2}\, n - 6\sqrt{n}$. For $1 \leqslant k \leqslant m^2$, let α_k denote the number of those indeterminates x_{ij} for which x_{ij} and $-x_{ij}$ together occur exactly k times in f. Then

$$\alpha_1 + 2\alpha_2 + 3\alpha_3 + \cdots \leqslant m^2,$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \cdots = n^2,$$

$$2n^2 - \alpha_1 \leqslant m^2,$$

and we will prove that α_1 is small.

We define a *line* of f to be either a row or a column of f, and let $X \subseteq \mathbf{x}$ denote the set of indeterminates occurring exactly once in f. If $x_{ij}, x_{kl} \in X$ occur (with a $\pm$ sign) in the same line of f, then either $i = k$ or $j = l$. Hence we can label those lines of f which contain at least two elements of X by "r" or "c"; a line receives the label r if it contains indeterminates from a single row of x, and c if from a single column of x.

Let G be the undirected bipartite graph whose vertices are the $2m$ lines of f, and with an edge between those lines that have an element of X at their intersection. Let e denote the number of edges of G. Any $x_{ij} \in X$ either

corresponds to an edge in G, or lies on an unlabelled line. There are at most $2m$ unlabelled lines, with at most one x_{ij} per line. Thus $\alpha_1 \leqslant 2m + e$.

We split G into four subgraphs, each of which has the same set of vertices as G, and the (disjoint) union of their edges is the edge set of G. G_1 is the graph whose edges connect vertices labelled r, G_2 is the graph connecting vertices labelled c, G_3 contains the edges which connect rows labelled r with columns labelled c, and G_4 contains the edges connecting rows labelled c with columns labelled r.

We first consider a connected component H of G_1 or G_2. Since all edges in H represent members of the same line of x, H contains at most n edges. Thus if H has at least $\sqrt{n}$ vertices, the average degree of vertices is at most $\sqrt{n}$. The same is trivially true for H with less than $\sqrt{n}$ vertices, and so the average degree in G_1 and G_2 is at most $\sqrt{n}$.

A four-cycle in G_3 or G_4 corresponds to a free square in f. By Lemma 4.3, G_3 and G_4 contain no four-cycles, and thus each have at most $(2m)^{3/2}$ edges (see Lovász (1979), Ch. 10, Ex. 36).

Assuming that $m \leqslant \sqrt{2}\, n - 6\sqrt{n} \leqslant \sqrt{2}\, n$, we find

$$e \leqslant 2 \cdot m \cdot \sqrt{n} + 2 \cdot (2m)^{3/2} \leqslant an^{3/2},$$

with $a = 2\sqrt{2} + 8\sqrt[4]{2}$. A simple calculation shows that

$$2n^2 - (2\sqrt{2}\, n + an^{3/2}) \leqslant 2n^2 - (2m + e) \leqslant 2n^2 - \alpha_1 \leqslant m^2 \leqslant (\sqrt{2}\, n - 6\sqrt{n})^2$$

implies $n \leqslant 70$. However, $\sqrt{2}\, n - 6\sqrt{n} \leqslant n < p_{\pm}(n)$ for $n \leqslant 70$, using Szegö (1913) (or Theorem 3.3). ∎

5. CONCLUSION AND OPEN QUESTIONS

One of the goals of Valiant's theory of "p-completeness" is to provide an analogue of the notorious Boolean conjecture P ≠ NP in a more structured setting (here: arithmetic computations), where powerful tools are available. The present criterion on maps relating the permanent and determinant yields an easy proof of (a generalization of) a theorem of Marcus and Minc. The main result is that $p(n) \geqslant p_{\pm}(n) > \sqrt{2}\, n - 6\sqrt{n}$. Meshulam (1987) has extended this to affine linear projections. It "remains" to improve this lower bounds, ultimately to superpolynomial in n. Theorem 3.1 by itself will not lead much further, since there are matrices $f \in (F \cup \{x_{11}, x_{12}, \ldots, x_{nn}\})^{m\times m}$ with $m \leqslant \sqrt{2}\, n + 2$, $\dim(\operatorname{im} f) = n^2$, and $\operatorname{im} f \cap \operatorname{sing} D_m = \emptyset$ (e.g. an upper

triangular matrix with ones on the diagonal and entries x_{ij} in the upper triangle).

From a geometric point of view, it would be interesting to determine the number of irreducible components of $\operatorname{sing} P_n$, and their dimensions. It is, however, not clear that such a result would yield new information on $p(n)$.

Many thanks go to László Babai and Ákos Seress for permission to include their proof. I have had many stimulating discussions with Volker Strassen and Michael Clausen about the subject and thank Allan Donsig and Gaston Gonnet for help in calculating $\operatorname{sing}(P_4)$ *on the Maple system.*

REFERENCES

Babai, L. and Seress, Á. 1987. Private communication, March 1987.

von zur Gathen, J. 1987. Feasible arithmetic computations: Valiant's hypothesis, *J. Symbolic Comput.* 3, to appear.

Lovász, L. 1979. *Combinatorial Problems and Exercises*, North-Holland.

Marcus, M. and Minc, H. 1961. On the relation between the determinant and the permanent, *Illinois J. Math.* 5:376–381.

Meshulam, R. 1987. Private communication, January 1987.

Minc, H. 1978. *Permanents*, Encyclopedia of Mathematics and its Applications 6, Addison-Wesley, Reading, Mass.

Pólya, G. 1913. Aufgabe 424, *Arch. Math. Phys.* 20:271.

Shafarevich, I. R. 1974. *Basic Algebraic Geometry*, Grundlehren Band 213, Springer.

Szegő, G. 1913. Zu Aufgabe 424, *Arch. Math. Phys.* 21:291–292.

Valiant, L. G. 1979a. Completeness classes in algebra, in *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, Atlanta Ga., pp. 249–261.

Valiant, L. G. 1979b. The complexity of computing the permanent, *Theoret. Comput. Sci.* 8:189–201.

van der Waerden, B. L. 1970. *Modern Algebra*, Vol. 1. Ungar, New York.

Received 4 March 1986; revised 3 December 1986

On Lewin and Vitek's Conjecture about the Exponent Set of Primitive Matrices

Zhang Ke Min
Department of Mathematics
Nanjing University
Nanjing, People's Republic of China

Submitted by Richard A. Brualdi

ABSTRACT

M. Lewin and Y. Vitek conjecture that every integer $\leqslant [\frac{1}{2}w_n]+1 = [\frac{1}{2}(n^2-2n+2)]+1$ is the exponent of some $n \times n$ primitive matrix. In this paper we prove that this conjecture is true except for $n = 11$. The problem of determining the exponent set E_n is completely solved.

INTRODUCTION

An $n \times n$ nonnegative square matrix $A = (a_{ij})$ is primitive if $A^k > 0$ for some positive integer k. The least such k is called the exponent of A and is denoted by $\gamma(A)$.

1. THE MAIN RESULT

In 1950, H. Wielandt [6] first stated the exact general upper bound for $\gamma(A)$, that is, $\gamma(A) \leqslant W_n = (n-1)^2+1$ for all $n \times n$ primitive matrices. In 1964, A. L. Dulmage and N. S. Mendelsohn [1] revealed the so-called *gaps* in the exponent set of $n \times n$ primitive matrices. Each gap is a set S of consecutive integers below W_n such that no $n \times n$ matrix A has an exponent in S. In 1981, M. Lewin and Y. Vitek [4] found the general method for determining all gaps between $[\frac{1}{2}W_n]+1$ and W_n, where $[x]$ denotes the greatest integer $\leqslant x$. And they conjectured that there are no gaps below