Some Computational Problems in Linear Algebra as hard as Matrix Multiplication[1]

Peter Bürgisser
Marek Karpinski
Thomas Lickteig

TR–91–005
January 14, 1991

[1]Submitted to *Computational Complexity*, Birkhäuser.

Some computational problems in linear algebra as hard as matrix multiplication

Peter Bürgisser* Marek Karpinski† Thomas Lickteig‡

Abstract

We define the complexity of a computational problem given by a relation using the model of computation trees together with the Ostrowski complexity measure. Natural examples from linear algebra are:

- KER_n: Compute a basis of the kernel for a given $n \times n$—matrix,
- OGB_n: Find an invertible matrix that transforms a given symmetric $n \times n$—matrix (quadratic form) to diagonal form,
- SPR_n: Find a sparse representation of a given $n \times n$—matrix.

To such a sequence of problems we assign an exponent similar as for matrix multiplication. For the complexity of the above problems we prove relative lower bounds of the form $aM_n - b$ and absolute lower bounds dn^2, where M_n denotes the complexity of matrix multiplication and a, b, d are suitably chosen constants. We show that the exponents of the problem sequences *KER, OGB, SPR* are the same as the exponent ω of matrix multiplication.

Key words: Problems, computation trees, straight line programs, Ostrowski complexity, derivations, matrix multiplication.

AMS(MOS) subject classifications: 68C20, 68C25.

1 Introduction

It is well known that matrix multiplication is crucial for many computational problems in linear algebra. Problems like matrix inversion, computation of the determinant or of all coefficients of the characteristic polynomial, LR-decomposition, and over the complex numbers also QR-decomposition and unitary transformation to Hessenberg form, are all known to be as hard as matrix multiplication.(See [3, 5, 8, 9, 13, 14, 17].) In this paper we study some computational problems in linear algebra that are specified more generally by a relation rather than by a function.

*International Computer Science Institute, 1947 Center Street, Suite 600, Berkeley, CA 94704.

†Institut für Informatik, Universität Bonn, Römerstr. 164, D-5300 Bonn 1, and International Computer Science Institute, 1947 Center Street, Suite 600, Berkeley, CA 94704.

‡International Computer Science Institute, 1947 Center Street, Suite 600, Berkeley, CA 94704. Research sponsored in part by the Institut für Informatik, Universität Bonn.

Let F denote a field of characteristic zero or an ordered field. The reader may keep in mind the two important examples $F = \mathbf{C}$ or $F = \mathbf{R}$. A problem is given by a relation

$$\Pi \subset F^m \times F^n.$$

(A relational connection between inputs and outputs is the natural way computational problems are specified; see also [4, 12].) Given an input $x \in F^m$ we are asked to find a $y \in F^n$ such that $(x, y) \in \Pi$. We say that a function

$$f : F^m \longrightarrow F^n$$

solves the problem Π if and only if

$$graph(f) \subset \Pi.$$

In order to investigate the complexity of a problem we use the model of a computation tree T using the operation symbols $F \sqcup \{0, 1, +, -, *, /\}$ (multiplications by scalars $\lambda \in F$ included) and the relation symbol $=$ (and $\leq$ when we are working over an ordered field). We define the cost of a computation tree T as the maximum number of multiplications and divisions T performes given an arbitrary input vector. (Compare [12, 15, 16, 18].) The complexity $C(f)$ of a function is then defined as the minimum cost of a tree computing f, and finally we put

$$C(\Pi) := min\{C(f) : f \text{ function solving } \Pi\}$$

for the complexity of the problem Π. In this paper we focus on the Ostrowski complexity measure which provides enough flexibility to carry through lower bound proofs. However, the upper bounds given in this paper also hold when all operations and comparisons are counted.

One of the leading problems in computational linear algebra is matrix multiplication. In our formal framework

$$MAMU_{(e,h,l)} := \{((A, B), C) \in (F^{e \times h} \times F^{h \times l}) \times F^{e \times l} : AB = C\}$$

Trivially

$$C(MAMU_{(ee',hh',ll')}) \leq C(MAMU_{(e,h,l)})e'h'l'. \tag{1}$$

We put

$$M_n := C(MAMU_{(n,n,n)}).$$

As a lower bound only the estimate $M_n \geq 2n^2 - 1$ is known ([2, 11]). The asymptotic behaviour of the sequence (M_n) is measured by the so-called exponent ω of matrix multiplication

$$\omega := inf\{\tau \in \mathbf{R} : M_n = O(n^\tau)\}.$$

(It is a well known fact that counting all arithmetic operations leads to the same asymptotic exponent.) The currently best known estimate is $2 \leq \omega < 2.376$ ([6, 19]).

We will study the following sequences of problems:

(1) 3-COMPRESSION:

$$3\text{-}CPR_n := \{((A_1, A_2, A_3), (B_1, B_2)) \in (F^{n \times n})^3 \times (F^{n \times n})^2 : A_1A_2A_3 = B_1B_2\}.$$

The investigation of this problem is motivated by the phenomenon that a corresponding problem for the addition of bitnumbers allows savings within the parallel boolean model. The task

given bitnumbers a, b, c find bitnumbers u, v such that $a + b + c = u + v$

can be solved more efficiently than by just adding up the numbers a, b, c using the so called *carry save adders*. (Cf. [20].) The lower bound we are going to prove shows that such a speed up is not possible for matrix multiplication (and sequential algorithms).

(2) KERNEL:

$$KER_n := \{(A,B) \in F^{n\times n} \times \bigsqcup_{i=0}^{n} F^{n\times i} : B \in F^{n\times(n-rk(A))}, rk(A)+rk(B)=n, AB=0\}.$$

This is of course the problem of computing a basis of the kernel for a given matrix.

(3) ORTHOGONAL BASIS:

$$OGB_n := \{(A,S) \in F^{n\times n} \times Gl_n : A \text{ symmetric, } SAS^T \text{diagonal}\}.$$

(4) SPARSE REPRESENTATION:

$$SPR_n := \{(A,(S,T,B)) \in F^{n\times n} \times (Gl_n^2 \times F^{n\times n}) : B = SAT, |supp(B)| \leq cn\}$$

($c \geq 1$ a fixed constant).

(5) SPARSENESS TRANSFORMATION MATRICES:

$$SPTM_n := \{(A,(S,T)) \in F^{n\times n} \times Gl_n^2 : |supp(SAT)| \leq cn\}$$

($c \geq 1$ a fixed constant).
In contrast to the SPARSE REPRESENTATION problem only the transformation matrices, but not the sparse representation matrix needs to be computed.

The main goal of this paper is to prove lower bounds on the complexity of the problems cited above in terms of the complexity of matrix multiplication. The proofs rest on differential methods from [17] and [3]. Section 2 contains the definitions of the formal framework and the model of computation.

Let us summarize our results: We can assign to any sequence $\Pi = (\Pi_n)$ of problems an exponent

$$\omega_\Pi := inf\{\tau \in \mathbf{R} : C(\Pi_n) = O(n^\tau)\}.$$

For any of the problem sequences Π listed under (1)–(5) we have

$$\omega_\Pi \leq \omega$$

(section 4) which follows easily from the recursive techniques given in [5, 13, 14]. In section 5 we prove for the sequences Π listed under (1)–(4) the lower bound

$$\forall n \;\; C(\Pi_n) \geq aM_n - bn^2$$

for suitably chosen constants $a, b > 0$. This implies immediately

$$\omega_\Pi \geq \omega,$$

provided that $\omega > 2$. For the sequence *SPTM* this estimate is also shown to be true.
The aim of section 6 is to remove this assumption " $\omega > 2$ " by showing absolute lower bounds

$$\forall n \;\; C(\Pi_n) \geq dn^2$$

(d a positive constant). The proof employs the notion of dimension for an affine variety. So for any of the sequences listed under (1)–(5) we have

$$\omega_\Pi = \omega.$$

2 Some terminology

We treat two cases in parallel. In the first case F denotes a field of characteristic zero, in the second F stands for an ordered field. (Think of the two examples $F = \mathbf{C}$ and $F = \mathbf{R}$.)

A *problem* Π is defined as being a subset

$$\Pi \subset F^m \times F^n.$$

We call $def(\Pi) := proj_{F^m}(\Pi)$ its domain of definition. We say that a partial function

$$f : F^m \supset def(f) \longrightarrow F^n$$

solves the problem Π if and only if

$$graph(f) \subset \Pi \text{ and } def(f) = def(\Pi).$$

In order to investigate these objects from the point of view of computations, we use the model of a computation tree. Let us shortly describe this notion; for a detailed discussion see [12, 15, 16, 18].

As the set Ω of operational symbols and the set P of relational symbols (together with arity functions) we take

$$\Omega = F \sqcup \{0, 1, +, -, *, /\}$$

and

$$\mathrm{P} = \{=\}$$

$$(\text{or } \mathrm{P} = \{=, \leq\} \text{ in case of an ordered field } (F, \leq)).$$

Let $s_1, s_2, \ldots$ be variables denoting storage locations in a computer. A *computation tree* T of type (Ω, P) with output length n is a binary tree together with a function that assigns

- to any simple vertex an operational instruction of the form

  $$s_i := \omega(s_{j_1}, \ldots, s_{j_k})$$

  where $k \geq 0$, $i, j_1, \ldots, j_k > 0$ and $\omega \in \Omega$ k-ary,

- to any branching vertex a test instruction of the form

  $$\rho(s_{j_1}, \ldots, s_{j_k})$$

  where $k \geq 0$, $j_1, \ldots, j_k > 0$ and $\rho \in \mathrm{P}$ k-ary,

- to any leaf an output instruction of the form

  $$(s_{j_1}, \ldots, s_{j_n})$$

  where $j_1, \ldots, j_n > 0$.

The assumption that all output lists have the same length n is made in order to simplify notation and is not essential. When fixing additionally an input length m such a computation tree T *computes* a partial function

$$f : F^m \supset def(f) \longrightarrow F^n$$

in the following way: given $\xi \in F^m$ we assign to the variables at the root of the tree the values $(\xi_1, \ldots, \xi_m, \infty, \infty, \ldots)$ and execute the instructions of T. This determines a directed path T_ξ from the root of T to a leaf or to a vertex with unexecutable instruction. We say $\xi \in def(f)$ if T_ξ ends up with a leaf. If this is the case the value of the output is $(f_1(\xi), \ldots, f_n(\xi))$. It is easy to see that for a directed path π from the root to the leaf the set

$$D_\pi := \{\xi \in F^m : T_\xi = \pi\} \subset F^m$$

is a locally closed subset and that the restriction of f to D_π is restriction of some rational function.

Now we are going to define the complexity of problems and functions. Let a cost function $c : \Omega \sqcup \mathrm{P} \to \mathbf{N}$ be given and T be a computation tree. By adding the costs along each path from the root to a leaf of the tree T and maximizing over all path we get the cost $cost_c(T)$ of T with respect to the cost function c. As the *complexity* $C_c(f)$ of a partial function

$$f : F^m \supset def(f) \longrightarrow F^n$$

we then define

$$C_c(f) := min\{cost_c(T) : T \text{ computation tree computing } f\},$$

and finally we call

$$C_c(\Pi) := min\{C_c(f) : f \text{ a function solving } \Pi\}$$

the *complexity* of the problem $\Pi \subset F^m \times F^n$ with respect to c. In the following we will assume that the cost function c is *arithmetic*, i.e. $c\,|_{\mathrm{P}} = 0$. (Notationally we will not distinguish between c and its restriction to Ω.)

Let A be a $F-$algebra, $f_1, \ldots, f_n \in A$, $I \in A^m$. The *complexity*

$$L_{A,c}(f_1, \ldots, f_n \mid I)$$

of $f_1, \ldots, f_n$ with respect to the input I and cost function $c : \Omega \to \mathbf{N}$ is the minimum cost of a $\Omega-$straight line program computing $f_1, .., f_n$ from the input I. Usually A is a localization of a quotient of the polynomial algebra $F[x_1, \ldots, x_m]$, and I will be chosen as the image of $(x_1, \ldots, x_m)$. In this case the input I will be notationally suppressed.

We mention that for a morphism $\psi : A \to A'$ of $F-$algebras we have

$$L_{A',c}(\psi(f_1), \ldots, \psi(f_n) \mid \psi(I)) \leq L_{A,c}(f_1, \ldots, f_n \mid I). \tag{2}$$

We outline our method for giving lower bounds on the complexity of problems. We say that a problem $\Pi \subset F^m \times F^n$ is *irreducibly defined* if $def(\Pi)$ is irreducible in the Zariski topology. Π is called *total* if $def(\Pi) = F^m$. A total problem is obviously irreducibly defined. The problems we will consider later on are all total. Now let an irreducibly defined problem $\Pi \subset F^m \times F^n$ be given and let $K := F(def(\Pi))$ denote the quotient field of the ring $F[def(\Pi)]$ of polynomials on $def(\Pi)$. Let $f : F^m \supset def(f) \to F^n$ be an optimal function solving Π and T be an optimal computation tree computing f, so

$$C_c(\Pi) = C_c(f) = cost_c(T).$$

We have a finite disjoint union

$$def(\Pi) = \bigcup\{D_\pi : \pi \text{ directed path from the root to a leaf}\}.$$

There must be a path π_0 such that D_{π_0} is Zariski dense in $def(\Pi)$ because $def(\Pi)$ is irreducible in the Zariski topology. Let us call such a path π_0 a typical one. If we are considering a field

F without ordering, there is exactly one typical path, because in this case a Zariski dense locally closed subset is open in $def(\Pi)$ and two of them must intersect. However, if we are working over an ordered field $(F, \leq)$ there might be many typical paths. Let π_0 be a typical one. Then we can consider the $g_i := f_i \mid_{D_{\pi_0}}$ as elements of the function field K and we easily see that

$$cost_c(T) \geq cost_c(\pi_0) \geq L_{K,c}(g_1, \ldots, g_n)$$

and

$$\forall \xi \in D_{\pi_0} \ \ (\xi, (g_1(\xi), \ldots, g_n(\xi)) \in \Pi.$$

So we have proved the following "arithmetic lemma"(compare [15, 16]):

Lemma 1 *Let $c : \Omega \sqcup \mathrm{P} \to \mathbf{N}$ be an arithmetic cost function, $\Pi \subset F^m \times F^n$ be a an irreducibly defined problem and $K = F(def(\Pi))$. Then there are elements $g_1, \ldots, g_n \in K$ such that*

$$(\xi, (g_1(\xi), \ldots, g_n(\xi))) \in \Pi \text{ for all } \xi \text{ in some Zariski dense subset of } def(\Pi)$$

and

$$C_c(\Pi) \geq L_{K,c}(g_1, \ldots, g_n).$$

This lemma gives a lower bound for the complexity of a total (irreducibly defined) problem in terms of the complexity of rational functions (algebraic functions) from which we only know that they satisfy certain relations. The clue will be to exploit this information in concrete cases.

In the following we will work exclusively with the Ostrowski cost function $c : \Omega \sqcup \mathrm{P} \to \mathbf{N}$ which is defined by $c(*) = c(/) = 1$, $c(\omega) = 0$ for all $\omega \in (\Omega \setminus \{*, /\}) \sqcup \mathrm{P}$. The index c will therefore be omitted.

3 Ostrowski complexity and differential methods

Let us recall some results from [17] and [3] on Ostrowski complexity of rational functions. Let $\lambda \in F^m$ and consider the local ring

$$\mathcal{O}_\lambda := \{ f \in F(x_1, \ldots, x_m) : \ f \text{ defined at } \lambda \ \}.$$

Lemma 2 *For given rational functions $f_1, \ldots, f_n \in F(x_1, \ldots, x_m)$ the equality*

$$L_{F(\underline{x})}(f_1, \ldots, f_n) = L_{\mathcal{O}_\lambda}(f_1, \ldots, f_n)$$

holds for Zariski-almost all $\lambda \in F^m$.

We omit the trivial proof.

It is well known that the image of an element $f \in \mathcal{O}_\lambda$ under the canonical imbedding

$$\mathcal{O}_\lambda \hookrightarrow F[\![y_1, \ldots, y_m]\!], \ \ x_i - \lambda \mapsto y_i$$

is the Taylor expansion $\sum_{k=0}^{\infty} f^{(k)}(\underline{y})$ in the point λ. (Here $f^{(k)}$ denotes the homogenous part of degree k of a polynomial $f \in F[y_1, \ldots, y_m]$.)

The next theorem will be used throughout in the paper.

Theorem 1 (**[17]**) *Let* $\lambda \in F^m$, $f_1, \ldots, f_n \in \mathcal{O}_\lambda$, $d \in \mathbf{N}$. *Then*

$$L_{F[y_1,\ldots,y_m]}(\{f_i^{(k)}(\underline{y}) : 0 \leq k \leq d,\ 1 \leq i \leq n\}) \leq \frac{d(d-1)}{2} L_{\mathcal{O}_\lambda}(f_1, \ldots, f_n),$$

where $\sum_{k=0}^{\infty} f_i^{(k)}(\underline{y})$ *denotes the Taylor expansion of* f_i *in the point* λ.

Observe that the complexity on the left-hand side is defined with respect to the polynomial ring $F[y_1, \ldots, y_m]$. Theorem 1 together with Lemma 2 for polynomials is Strassen's result on "Vermeidung von Divisionen". A well known consequence of which is

$$M_n = C(MAMU_{(n,n,n)}) = L_{F[\underline{X},\underline{Y}]}(\{\sum_{l=1}^{n} X_{il}Y_{lj} : 1 \leq i,j \leq n\}).$$

Our second main tool is the famous "Derivation Theorem".

Theorem 2 (**[3]**) *Let* $f \in F(x_1, \ldots, x_m)$ *be a rational function. Then*

$$L_{F(\underline{x})}(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_m}) \leq 3L_{F(\underline{x})}(f).$$

4 Relative upper bounds

We recall the definitions of the computational problems we are interested in.

- t–COMPRESSION:

  $$t\text{-}CPR_n := \{((A_1, \ldots, A_t), (B_1, \ldots, B_{t-1})) \in (F^{n\times n})^t \times (F^{n\times n})^{t-1} : A_1A_2 \cdots A_t = B_1B_2 \cdots B_{t-1}\}.$$

  ($t \geq 2$ a natural number). Obviously

  $$C((t+1)\text{-}CPR_n) \leq C(t\text{-}CPR_n),\ C(2\text{-}CPR_n) = MAMU_{(n,n,n)}.$$

- KERNEL:

  $$KER_n := \{(A,B) \in F^{n\times n} \times \bigsqcup_{i=0}^{n} F^{n\times i} : B \in F^{n\times(n-rk(A))}, rk(A) + rk(B) = n, AB = 0\}.$$

- ORTHOGONAL BASIS:

  $$OGB_n := \{(A,S) \in F^{n\times n} \times Gl_n : A \text{ symmetric}, SAS^T \text{diagonal}\}.$$

- SPARSE REPRESENTATION:

  $$SPR_n := \{(A,(S,T,B)) \in F^{n\times n} \times (Gl_n^2 \times F^{n\times n}) : B = SAT, |supp(B)| \leq cn\}$$

  ($c \geq 1$ a fixed "sparseness" constant).

- SPARSENESS TRANSFORMATION MATRICES:

$$SPTM_n := \{(A,(S,T)) \in F^{n\times n} \times Gl_n^2 : |supp(SAT)| \leq cn\}$$

($c \geq 1$ a fixed constant).

The following theorem gives an upper bound relative to the complexity of matrix multiplication.

Theorem 3 *The exponent for any of the sequences of problems*

$$t\text{-}CPR, KER, OGB, SPR, SPTM$$

is less or equal the exponent ω of matrix multiplication.

The proof is based on ideas from [5, 13, 14]. See also [1, pages 233–240] and [9]. The proceeding is to subdivide the occuring matrices into blocks, to perform a sort of Gaussian Elimination blockwise using a fast hypothetical matrix multiplication algorithm, and then to continue recursively. We leave the details to the reader. For the problem *t-CPR* the statement is of course trivial.

Remark: Theorem 3 remains true when we count all rational operations and tests at unit cost.

5 Relative lower bounds

We are going to prove lower bounds in terms of M_n for the various problems defined above.

Theorem 4 *The sequence 3-CPR satisfies*

$$C(3\text{-}CPR_n) \geq \frac{1}{3}M_n - n^2.$$

Proof: Let A, B, C be $n \times n$—matrices whose entries are indeterminates over F and put $K := F(A_{ij}, B_{ij}, C_{ij})$. By Lemma 1 there are $U, V \in K^{n\times n}$ such that

$$UV = ABC$$

and

$$L_K(U,V) \leq C(3\text{-}CPR_n).$$

If we take into consideration that the trace of the product of two $n \times n$—matrices can be computed with n^2 multiplications, we get

$$L_K(Tr(ABC)) \leq C(3\text{-}CPR_n) + n^2.$$

Furthermore

$$\frac{\partial Tr(ABC)}{\partial A_{ij}} = (BC)_{ji}.$$

Theorem 2 implies now

$$M_n = L_K(BC) \leq 3C(3\text{-}CPR_n) + 3n^2,$$

which completes the proof of the theorem. □

Remark: We conjecture that a similar lower bound holds for the problem sequences *t-CPR* when $t > 3$.

Theorem 5 *The sequence KER satisfies*

$$C(KER_n) \geq M_{\lfloor n/4 \rfloor}.$$

Proof: W.l.o.g. we may assume that $n = 4m$, $m \in \mathbf{N}$. Let X, Y denote $2m \times 2m$-matrices whose entries are indeterminates over F. We put $K := F(X_{ij}, Y_{ij})$ and $R := F[X_{ij}, Y_{ij}]$. When we apply Lemma 1 to the restricted problem

$$KER_n \cap \left(\left\{ \begin{pmatrix} \xi & \eta \\ 0 & 0 \end{pmatrix} : \xi, \eta \in F^{2m \times 2m} \right\} \times \bigsqcup_{i=0}^{n} F^{n \times i} \right),$$

we see that there exists a matrix $B \in K^{4m \times 2m}$ satisfying

$$rk(B) = 2m,\ (X, Y)B = 0$$

and

$$L_K(B) \leq C(KER_n).$$

There must be a matrix $U \in Gl_{2m}(K)$ such that

$$B = \begin{pmatrix} X^{-1}U \\ Y^{-1}U \end{pmatrix}.$$

We therefore have

$$L_K(X^{-1}U, Y^{-1}U) \leq C(KER_n).$$

From Lemma 2 we obtain that there are $\xi, \eta \in Gl_{2m}(F)$ such that

$$X, Y, U \in Gl_{2m}(\mathcal{O}_{(\xi,\eta)}) \text{ and } L_{\mathcal{O}_{(\xi,\eta)}}(X^{-1}U, Y^{-1}U) = L_K(X^{-1}U, Y^{-1}U).$$

Thanks to the fact that we do not count linear operations we may replace U by $UU(\xi, \eta)^{-1}$ and therefore assume that $U(\xi, \eta) = E$ (E = identity matrix). Application of the isomorphism

$$\varphi : \mathcal{O}_{(\xi,\eta)} \to \mathcal{O}_{(E,E)}, \varphi(X) := X\xi, \varphi(Y) := Y\eta$$

shows that we may assume w.l.o.g. that $\xi = \eta = E$. (Cf. (2).) Furthermore

$$L_{\mathcal{O}_{(0,0)}}((E - X)^{-1}V, (E - Y)^{-1}V) = L_{\mathcal{O}_{(E,E)}}(X^{-1}U, Y^{-1}U),$$

when we put $V := U(E - X, E - Y)$. We use now Theorem 1 with $d = 2$ and get

$$L_R(X^2 + XV^{(1)} + V^{(2)}, Y^2 + YV^{(1)} + V^{(2)}) \leq L_{\mathcal{O}_{(0,0)}}((E - X)^{-1}V, (E - Y)^{-1}V),$$

where $V = E + V^{(1)} + V^{(2)} + \cdots$ denotes the Taylor expansion of V in the point $(0, 0)$. The complexity on the left-hand side can be estimated from below by

$$L_R(X^2 - Y^2 + (X - Y)V^{(1)}).$$

We write

$$X = \begin{pmatrix} X^{11} & X^{12} \\ X^{21} & X^{22} \end{pmatrix}, Y = \begin{pmatrix} Y^{11} & Y^{12} \\ Y^{21} & Y^{22} \end{pmatrix}$$

where $X^{ij}, Y^{ij} \in R^{m\times m}$ and make the linear substitution

$$\psi : R \to R, \psi(X) := \begin{pmatrix} 0 & X^{12} \\ X^{21} & X^{22} \end{pmatrix}, \psi(Y) := \begin{pmatrix} 0 & X^{12} \\ 0 & 0 \end{pmatrix}.$$

One calculates immediately that

$$\psi(X^2 - Y^2 + (X-Y)V^{(1)}) = \begin{pmatrix} X^{12}X^{21} & X^{12}X^{22} \\ P & Q \end{pmatrix}$$

for some $P, Q \in R^{m\times m}$. Therefore

$$M_m = L_R(X^{12}X^{21}) \leq C(KER_n).$$

□

Theorem 6 *The sequence OGB satisfies*

$$C(OGB_n) \geq \frac{1}{3}M_{\lfloor n/4 \rfloor} - 4n^2 - n.$$

Proof W.l.o.g. we may assume that $n = 4m$, $m \in \mathbf{N}$. Let A denote a symmetric $n \times n$—matrix whose entries are indeterminates over F. Put $K := F(A_{ij} : i \leq j \leq n)$ and $R := F[A_{ij} : i \leq j \leq n]$. By Lemma 1 there is a matrix $S \in Gl_n(K)$ such that

$$D := SAS^T \text{ is diagonal}$$

and

$$L_K(S) \leq C(OGB_n).$$

By writing

$$\begin{bmatrix} D_{11} \\ \vdots \\ D_{nn} \end{bmatrix} = D \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = S(A(S^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix})),$$

we see that D can be computed from A and S with $2n^2$ multiplications. We have

$$Tr(A^{-1}) = Tr(S^T(D^{-1}S)).$$

Therefore

$$L_K(Tr(A^{-1}) \leq L_K(S) + 4n^2 + n.$$

We proceed now similar as in [8]. Let $V \in F^{n\times n}$ be symmetric and ϵ be an indeterminate over K. Then

$$Tr((A+\epsilon V)^{-1}) = Tr(A^{-1}) + \epsilon \sum_{i\leq j} \frac{\partial Tr(A^{-1})}{\partial A_{ij}} V_{ij} + O(\epsilon^2).$$

On the other hand one easily calculates

$$Tr((A+\epsilon V)^{-1}) = Tr(A^{-1}) - \epsilon Tr(A^{-1}VA^{-1}) + O(\epsilon^2).$$

Comparing the two equations we get

$$\frac{\partial Tr(A^{-1})}{\partial A_{ij}} = \begin{cases} -2(A^{-2})_{ij} & \text{, if } i \neq j \\ -(A^{-2})_{ij} & \text{, otherwise.} \end{cases}$$

From the Derivation Theorem 2 we deduce

$$\frac{1}{3}L_K(A^{-2}) \leq L_K(Tr(A^{-1})) \leq L_K(S) + 4n^2 + n.$$

Because the matrices of the form $\mu\mu^T$ ($\mu \in F^{n\times n}$) are a Zariski dense part of the symmetric matrices in $F^{n\times n}$ we conclude by Lemma 2 that there exists a matrix $\mu \in Gl_n(F)$ such that

$$L_{\mathcal{O}_{\mu\mu^T}}(A^{-2}) = L_K(A^{-2}).$$

Furthermore using the $F-$algebra isomorphism $\mathcal{O}_{\mu\mu^T} \to \mathcal{O}_E, A \mapsto \mu A\mu^T$ we see

$$L_{\mathcal{O}_{\mu\mu^T}}(A^{-2}) = L_{\mathcal{O}_E}(A^{-2}) = L_{\mathcal{O}_0}((E-A)^{-2}).$$

By Theorem 1 with $d = 2$ and taking into account that

$$(E-A)^{-2} = E + 2A + 3A^2 + \ldots$$

we get

$$L_R(A^2) \leq L_{\mathcal{O}_0}((E-A)^{-2}).$$

We divide the matrix A into $m \times m-$blocks $A^{ij} \in R^{m\times m}$ and define the substitution

$$\psi : R \to R, \psi(A) := \begin{pmatrix} 0 & 0 & A^{13} & 0 \\ 0 & 0 & A^{23} & 0 \\ (A^{13})^T & (A^{23})^T & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Obviously

$$\psi(A)^2 = \begin{pmatrix} A^{13}(A^{13})^T & A^{13}(A^{23})^T & 0 & 0 \\ A^{23}(A^{13})^T & A^{23}(A^{23})^T & 0 & 0 \\ 0 & 0 & W & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where $W = (A^{13})^T A^{13} + (A^{23})^T A^{23}$. Hence

$$M_m = L_R(A^{13}(A^{23})^T) \leq L_R(A^2) \leq 3(C(OGB_n) + 4n^2 + n).$$

□

Throughout the following $c \geq 1$ will denote a fixed constant. We call a $n \times n$—matrix A *sparse* (with respect to the sparseness constant c) if

$$|supp(A)| \leq cn.$$

Observe that an arbitrary $n \times n$—matrix can be multiplied with a sparse matrix (from the left or right) using only cn^2 multiplications.

Theorem 7 *The sequence SPR satisfies*

$$C(SPR_n) \geq \frac{1}{9}M_{\lfloor n/3 \rfloor} - (2 + 10c/3)n^2.$$

Proof W.l.o.g. we may assume that $n = 3m$, $m \in \mathbf{N}$. Let A denote a $n \times n$–matrix whose entries are indeterminates over F. We set $K := F(A_{ij} : i,j \leq n)$ and $R := F[A_{ij} : i,j \leq n]$. By Lemma 1 there exists $\tilde{S}, \tilde{T} \in Gl_n(K)$ and a sparse matrix $\tilde{B} \in K^{n \times n}$ such that

$$\tilde{B} = \tilde{S}A\tilde{T}$$

and

$$L_K(\tilde{S}, \tilde{T}, \tilde{B}) \leq C(SPR_n).$$

By Lemma 2 we may choose a matrix $\alpha \in Gl_n(F)$ such that

$$\tilde{S}, \tilde{T}, \tilde{B} \in Gl_n(\mathcal{O}_\alpha) \text{ and } L_{\mathcal{O}_\alpha}(\tilde{S}, \tilde{T}, \tilde{B}) = L_K(\tilde{S}, \tilde{T}, \tilde{B}).$$

By applying the isomorphism $\mathcal{O}_\alpha \to \mathcal{O}_0, A \mapsto \alpha(E - A)$ we see that there exist $S, T, B \in Gl_n(\mathcal{O}_0)$, satisfying

$$B = S(E - A)T, \ B \text{ sparse}$$

and

$$L_{\mathcal{O}_0}(S, T, B) = L_{\mathcal{O}_\alpha}(\tilde{S}, \tilde{T}, \tilde{B}).$$

Let $S = S^{(0)} + S^{(1)} + \ldots$, $T = T^{(0)} + T^{(1)} + \ldots$, $B = B^{(0)} + B^{(1)} + \ldots$ denote the Taylor expansions in 0 of S, T, B respectively. The matrices $B^{(k)}$ are also sparse. Theorem 1 implies

$$L_R(S^{(2)}, S^{(3)}, T^{(2)}, T^{(3)}) \leq 3L_{\mathcal{O}_0}(S, T, B).$$

By comparing the third order terms in the Taylor expansion in 0 of both sides of the equation

$$(E - A)^{-1} = TB^{-1}S$$

we get

$$Tr(A^3) = \sum_{i+j+k=3} Tr(T^{(i)}((B^{-1})^{(j)}S^{(k)})).$$

We are going to show now that the products $(B^{-1})^{(j)}S^{(k)}$ $(j + k \leq 3)$ can be computed from $B^{(2)}, B^{(3)}, S^{(2)}, S^{(3)}$ with only $10cn^2$ multiplications:

A short calculation yields (put $\gamma := (B^{(0)})^{-1}$)

$$(B^{-1})^{(1)} = -\gamma B^{(1)}\gamma, \tag{3}$$

$$(B^{-1})^{(2)} = -\gamma B^{(2)}\gamma + \gamma B^{(1)}\gamma B^{(1)}\gamma, \tag{4}$$

$$(B^{-1})^{(3)} = -\gamma B^{(3)}\gamma + \gamma B^{(2)}\gamma B^{(1)}\gamma + \gamma B^{(1)}\gamma B^{(2)}\gamma + \gamma B^{(1)}\gamma B^{(1)}\gamma B^{(1)}\gamma. \tag{5}$$

Observe furthermore that a product

$$\gamma\Sigma_1\gamma\Sigma_2 \cdots \gamma\Sigma_t\gamma\Gamma,$$

where

$$\gamma \in F^{n \times n}, \ \Gamma \in R^{n \times n}$$

and

$$\Sigma_i \in R^{n \times n} \text{ sparse } (i = 1, \ldots, t),$$

can be computed from the matrices Σ_i, Γ with only ctn^2 nonscalar multiplications. (Compute from the righthand side to the left.) Taking this into account the upper bound $10cn^2$ follows now easily.

The result of this intermediate reasoning gives the upper bound

$$L_R(Tr(A^3)) \leq 3C(SPR_n) + 6n^2 + 10cn^2.$$

We subdivide A into $m \times m$–blocks $A^{ij} \in R^{m \times m}$ and make the substitution

$$\psi : R \to R, \psi(A) := \begin{pmatrix} 0 & A^{12} & 0 \\ 0 & 0 & A^{23} \\ A^{13} & 0 & 0 \end{pmatrix}.$$

One easily verifies

$$\psi(A)^3 = \begin{pmatrix} A^{12}A^{23}A^{31} & 0 & 0 \\ 0 & A^{23}A^{31}A^{12} & 0 \\ 0 & 0 & A^{31}A^{12}A^{23} \end{pmatrix}$$

and hence

$$\psi(Tr(A^3)) = Tr(\psi(A)^3) = 3Tr(A^{12}A^{23}A^{31}).$$

So we showed that

$$L_R(A^{12}A^{23}A^{31}) \leq L_R(Tr(A^3)).$$

When we apply the Derivation Theorem 2 we finally get

$$\frac{1}{3}M_m \leq L_R(A^{12}A^{23}A^{31})$$

which implies the desired bound

$$\frac{1}{9}M_m - (2 + 10c/3)n^2 \leq C(SPR_n).$$

□

As an immediate consequence of Theorem 3 and Theorems 4 - 7 we get the following

Corollary 1 *Any of the sequences of problems*

3-CPR, KER, OGB, SPR

has as exponent the exponent ω of matrix multiplication, provided that $\omega > 2$ in case of 3-CPR, OGB, SPR .

For the sequence *SPTM* we will only make a statement about the exponent. We need the following

Lemma 3 *The sequence of problems $MAMU_{(n,n,\lfloor\sqrt{n}\rfloor)}$ has an exponent strictly smaller than the exponent ω of matrix multiplication, provided that $\omega > 2$.*

Proof: We assume $\omega > 2$ and choose ϵ satisfying $0 < \epsilon < \omega/2 - 1$. For a suitable constant $d > 0$ and all squares n we have by inequality (1)

$$C(MAMU_{(n,n,\sqrt{n})}) \leq C(MAMU_{(\sqrt{n},\sqrt{n},\sqrt{n})})n \leq d(\sqrt{n})^{\omega+2\epsilon}n.$$

Therefore

$$C(MAMU_{(n,n,\sqrt{n})}) = O(n^{\omega/2+\epsilon+1}).$$

But $\omega/2 + \epsilon + 1 < \omega$ and the statement follows. □

Theorem 8 *The exponent for the sequence SPTM equals the exponent ω of matrix multiplication, provided that $\omega > 2$.*

Proof: Suppose $\omega > 2$. Since we already proved Theorem 7 it is sufficient to show the following:

Given $(A, S, T) \in F^{n\times n} \times Gl_n^2$ and the information that SAT is sparse, then we can compute SAT with cost $O(n^\tau)$, where $\tau < \omega$.

Put $B := SAT$ and assume B being sparse. For $i \in \{1, \ldots, n\}$ we define

$$I_i := \{j \in \{1, \ldots, n\} : B_{ij} \neq 0\}.$$

In order to simplify notation we may assume w.l.o.g. that $(|I_i|)_{j=1,\ldots,n}$ is an increasing sequence. We set

$$M := \max\{i : |I_i| \leq \lfloor\sqrt{n}\rfloor\}.$$

Then

$$n - M \leq c\sqrt{n}.$$

We now choose a matrix $\alpha \in F^{n\times\lfloor\sqrt{n}\rfloor}$ with the property that all its subdeterminants are different from zero. According to the preceeding lemma we can compute the product C

$$C := B\alpha = S(A(T\alpha))$$

with cost $O(n^\tau)$, where $\tau < \omega$. However, the first M rows of B can be computed from C without any nonlinear operations: for a fixed $i \leq M$ we have

$$\forall\ k \leq \lfloor\sqrt{n}\rfloor \quad \sum_{j\in I_i} B_{ij}\alpha_{jk} = C_{ik}.$$

As for all $i \leq M$ $(\alpha_{jk})_{j\in I_i, k\leq |I_i|} \in Gl_{|I_i|}(F)$ we get all B_{ij} $(i \leq M)$ from C with linear operations only.

In order to get the remaining $n - M$ rows of B we do simply the following. We choose a matrix $\beta \in Gl_{n-M}(F)$ and put

$$\gamma := (0, \beta) \in F^{(n-M)\times n}.$$

The product

$$\gamma B = ((\gamma S)A)T$$

can be computed with only $O(n^\tau)$ nonlinear operations as well. But from γB we can obtain $(B_{ij})_{i=M+1,\ldots,n,\ j=1,\ldots,n}$ with linear operations only. □

6 Absolute lower bounds

The aim of this section is to show that the assumption "$\omega > 2$" in Corollary 1 and Theorem 8 is unnecessary. We will do so by proving lower bounds of the type

$$\text{constant} \cdot n^2$$

for the various computational problems we considered before. There is no harm in assuming that F is algebraically closed.

We need the subsequent proposition

Proposition 1 *Let be* $\lambda \in F^m, f_1, \ldots, f_n \in \mathcal{O}_\lambda, f_i = \sum_{k=0}^{\infty} f_i^{(k)}$ *its Taylor expansion. Then there is a linear subspace* $S \subset F^m$ *contained in the closed cone*

$$Z := \{\eta \in F^m : f_i^{(k)}(\eta) = 0 \text{ for all } i \in \{1, \ldots, n\}, k \geq 2\}$$

such that

$$dim S \geq m - L_{\mathcal{O}_\lambda}(f_1, \ldots, f_n).$$

In particular,

$$dim Z \geq m - L_{\mathcal{O}_\lambda}(f_1, \ldots, f_n).$$

Proof: W.l.o.g. we may assume $\lambda = 0$. We make induction on $r := L_{\mathcal{O}_0}(f_1, \ldots, f_n)$. The start "$r = 0$" is trivial. Assume now that $r > 0$. It is easy to see that there is a hyperplane $H \subset F^m$ such that the complexity of the restricted functions $\tilde{f}_i := f_i \mid_H$ (considered as elements of the local ring of H in 0) is strictly less than r. The induction hypothesis implies the existence of a linear subspace $S \subset H$ satisfying

$$S \subset \{\eta \in H : f_i^{(k)}(\eta) = 0 \text{ for all } i \in \{1, \ldots, n\}, k \geq 2\}$$

and

$$dim S = (m-1) - (r-1) = m - r,$$

as asserted. □

Theorem 9 *The problems t-CPR*$_n$, *OGB*$_n$, *SPR*$_n$ *and SPTM*$_n$ *have the following lower bounds:*

$$\begin{array}{lcl} C(\text{t-CPR}_n) & \geq & \frac{1}{2}n^2(1 + 1/min\{t, n\}), \\ C(OGB_n) & \geq & n^2/18 - 17n/6, \\ C(SPR_n) & \geq & n^2/3 - 3cn \text{ (c=sparseness constant)}, \\ C(SPTM_n) & \geq & n^2/3 - 3cn. \end{array}$$

Proof: We give a detailed proof for the sequence *SPTM*, for the other problems we only sketch the way of argumentation.

SPTM: Let A denote a $n \times n$—matrix whose entries are indeterminates over F. The same reasoning as in the proof of Theorem 7 shows that there are $S, T, B \in Gl_n(\mathcal{O}_0)$ such that

$$B = S(E - A)T \text{ sparse},$$

and

$$L_{\mathcal{O}_0}(S, T) \leq C(SPTM_n).$$

By Proposition 1 the cone

$$Z := \{\alpha \in F^{n \times n} : S^{(k)}(\alpha) = T^{(k)}(\alpha) = 0 \text{ for all } k \geq 2\}$$

has codimension

$$codim(Z) \leq L_{\mathcal{O}_0}(S, T).$$

From $(E - A)^{-1} = TB^{-1}S$ we get

$$A^3 = \sum_{i+j+k=3} T^{(i)}(B^{-1})^{(j)}S^{(k)}. \tag{6}$$

By the Dimension Theorem [7, page 48]

$$Z' := Z \cap \{\alpha \in F^{n\times n} : B^{(k)}(\alpha) = 0 \text{ for } k = 1,2,3\ \}.$$

is a closed cone of codimension

$$codim_{F^{n\times n}}(Z') \leq L_{\mathcal{O}_0}(S,T) + 3cn,$$

since the matrices $B^{(k)}$ are sparse. From the equations (3)– (5) of section 5 it follows

$$Z' \subset \{\alpha \in F^{n\times n} : (B^{-1})^{(k)}(\alpha) = 0 \text{ for } k = 1,2,3\}.$$

By equation 6 above

$$Z' \subset \{\alpha \in F^{n\times n} : \alpha^3 = 0\}.$$

The subsequent Lemma 4 and a comparison of dimensions lead to the inequality

$$L_{\mathcal{O}_0}(S,T) \geq n^2/3 - 3cn$$

which proves the statement.

t-CPR: There are $\alpha \in (F^{n\times n})^t$, $B_1, \ldots, B_{t-1} \in \mathcal{O}_\alpha$ such that

$$A_1 \cdots A_t = B_1 \cdots B_{t-1} \text{ and } L_{\mathcal{O}_\alpha}(B_1, \ldots, B_{t-1}) \leq C(t\text{-}CPR_n).$$

We set

$$Z := \{\beta \in (F^{n\times n})^t : B_i^{(k)}(\beta) = 0 \text{ for all } i \leq t-1, k \geq 2.\}$$

Obviously

$$Z \subset \{\beta \in (F^{n\times n})^t : \beta_1 \cdots \beta_t = 0\} =: Z_{n,t}.$$

But by Proposition 1 we have

$$codim(Z) \leq L_{\mathcal{O}_\alpha}(B_1, \ldots, B_{t-1}) \leq C(t\text{-}CPR_n),$$

hence

$$codim(Z_{n,t}) \leq C(t\text{-}CPR_n).$$

The subsequent Lemma 4 implies now the desired statement.

OGB: We proceed as for the problems *SPTM*, but instead of Lemma 4 we use

$$dim(\{\alpha \in F^{n\times n} : \alpha \text{ symmetric, } \alpha^3 = 0\}) \leq 4n^2/9 + n/3,$$

which can be derived from the formula

$$dim(\{\alpha \in F^{n\times n} : \alpha \text{ symmetric, } rk(\alpha) \leq r\}) = rn - r(r-1)/2$$

if we take into account that

$$\forall \alpha \in F^{n\times n}\ (\alpha^3 = 0 \Longrightarrow rk(\alpha) \leq 2n/3).$$

SPR: The statement follows trivially from $C(SPR_n) \geq C(SPTM_n)$. □

Lemma 4 *The closed subvarieties*

$$N_{n,t} := \{\alpha \in F^{n\times n} : \alpha^t = 0\} \subset F^{n\times n},$$

$$Z_{n,t} := \{(\alpha_1, \ldots, \alpha_t) \in (F^{n\times n})^t : \alpha_1 \cdots \alpha_t = 0\} \subset (F^{n\times n})^t$$

have the codimensions

$$codim(N_{n,t}) = \sigma_{n,t},$$

$$codim(Z_{n,t}) = (n^2 + \sigma_{n,t})/2$$

where

$$\sigma_{n,t} := min\{\sum_{i=1}^{t} m_i^2 : (m_1, \ldots, m_t) \text{ partition of } n\}.$$

The sequence $\sigma_{n,t}$ satisfies

$$\begin{array}{ll} \sigma_{n,t} = n & \text{if } t \geq n, \\ \sigma_{n,t} \geq n^2/t & \text{if } t \leq n \text{ (with equality if } t|n). \end{array}$$

This lemma is classical and can be proved by standard techniques. For the reader's convenience we include a proof.

Proof:
For $N_{n,t}$: Let β be a matrix in Jordan normal form, i.e.

$$\beta = diag(J(n_1, \lambda_1), \ldots, J(n_s, \lambda_s)),$$

where $(n_1, \ldots, n_s)$ is a partition of n and $J(n_i, \lambda_i)$ is defined as

$$J(n_i, \lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{pmatrix} \in F^{n_i \times n_i}.$$

We denote the dimension of the conjugacy class of β by $d_\lambda(n_1, \ldots, n_s)$. It is easy to see that

$$\beta^t = 0 \iff \forall i \; \lambda_i = 0, n_i \leq t.$$

This implies at once

$$dim(N_{n,t}) = max\{d_0(n_1, \ldots, n_s) : (n_1, \ldots, n_t) \text{ partition of } n \text{ with } n_i \leq t \text{ for all } i\}.$$

The value of $d_0(n_1, \ldots, n_s)$ can be exactly determined, namely

$$d_0(n_1, \ldots, n_s) = n^2 - \sum_{j=1}^{t} m_j^2,$$

where $(m_1, \ldots, m_t)$ denotes the partition dual to $(n_1, \ldots, n_s)$. (See [10, page 192].) Using this fact we conclude immediately

$$codim(N_{n,t}) = min\{\sum_{j=1}^{t} m_j^2 : (m_1, \ldots, m_t) \text{ partition of } n\} = \sigma_{n,t}.$$

For $Z_{n,t}$: We assign to an element $\alpha \in (F^{n\times n})^t$ a sequence $(V_0, V_1, \ldots, V_t)$ of linear subspaces of F^n by putting

$$V_t := F^n, \;\; V_{i-1} := \alpha_i(V_i) \text{ for } i = t, \ldots, 1.$$

We set

$$d_i := dim(V_i), \;\; m_i := d_i - d_{i-1}$$

and call $(m_1, \ldots, m_t)$ the pattern of α. Obviously

$$\alpha \in Z_{n,t} \Longleftrightarrow \sum_{i=1}^{t} m_i = n.$$

Let now a pattern $m = (m_1, \ldots, m_t)$ satisfying the condition $\sum_{i=1}^{t} m_i = n$ be fixed. The subvariety

$$\{\alpha \in (F^{n\times n})^t : \; \alpha \text{ has pattern } m\} \subset (F^{n\times n})^t$$

is isomorphic to the subvariety

$$\Phi := \{((V_1, \ldots, V_t), \alpha) : V_t = F^n, \; V_{i-1} = \alpha_i(V_i), \; dim(V_i) = d_i \text{ for all } i\}$$

of $\prod_{i=1}^{t} Gr_{n,d_i} \times (F^{n\times n})^t$. ($Gr_{n,d_i}$ denote the Grassmann varieties.) It is well known that $dim(Gr_{n,d}) = d(n-d)$.

All fibres of the projection $\Phi \to \prod_{i=1}^{t} Gr_{n,d_i}$ have dimension $tn^2 - \sum_{i=1}^{t} d_i(n - d_{i-1})$. We get now easily

$$codim(\Phi) = \sum_{j=1}^{t} d_j(d_j - d_{j-1}) = \sum_{i\le j} m_i m_j,$$

hence

$$codim(\Phi) = (n^2 + \sum_{i=1}^{t} m_j^2)/2.$$

Therefore

$$codim(Z_{n,t}) = (n^2 + \sigma_{n,t})/2.$$

□

Putting all our information together we get the final result

Corollary 2 *Any of the sequences of problems*

$$3\text{-}CPR, KER, OGB, SPR, SPTM$$

has as exponent the exponent ω of matrix multiplication.

References

[1] *A.V. Aho*, *J.E. Hopcroft*, and *J.D. Ullman*, The design and analysis of computer algorithms, Reading MA: Addison–Wesley, 1974.

[2] *A. Alder* and *V. Strassen*, On the algorithmic complexity of associative algebras, Theor. Computer Science **15**(1981), 201–211.

[3] *W. Baur* and *V. Strassen*, The complexity of partial derivatives, Theor. Computer Science **22**(1982), 317–330.

[4] *L. Blum, M. Shub,* and *S. Smale*, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, Bull. Amer. Math. Soc. **21**(1989), 1–46.

[5] *J. Bunch* and *J. Hopcroft*, Triangular factorizationand inversion by fast matrix multiplication, Math. Comp. **28**(1974), 231–236.

[6] *D. Coppersmith* and *S. Winograd*, Matrix multiplication via arithmetic progressions, Proc. 19th ACM STOC, New York(1987) , 1–6.

[7] *R. Hartshorne*, Algebraic Geometry, Graduate Texts in Mathematics Vol. 52, Springer Verlag, 1977.

[8] *K. Kalorkoti*, The trace invariant and matrix inversion, Theor. Computer Science **59**(1988), 277–286.

[9] *W. Keller–Gehrig*, Fast algorithms for the characteristic polynomial, Theor. Computer Science **36**(1985), 309–317.

[10] *H. Kraft*, Geometric methods in representation theory, in: Representations of Algebras, Workshop Proc., Puebla, Mexico 1980, LNM **944**, Berlin–Heidelberg–New York 1982.

[11] *J.C. Lafon* and *S. Winograd*, A lower bound for the multiplicative complexity of the product of two matrices, (unpublished) manuscript, 1978.

[12] *T. Lickteig*, On semialgebraic decision complexity, Tech. Rep. TR-90-052 Int. Comp. Science Inst., Berkeley, and Univ. Tübingen, Habilitationsschrift, to appear.

[13] *A. Schönhage*, Unitäre Transformationen grosser Matrizen, Num. Math. **20**(1973), 409–417.

[14] *V. Strassen*, Gaussian elimination is not optimal, Numer. Mathematik **13**(1969), 354–356.

[15] *V. Strassen*, Berechnung und Programm I, Acta Informatica **1**(1973), 320–335.

[16] *V. Strassen*, Berechnung und Programm II, Acta Informatica **2**(1973), 64–79.

[17] *V. Strassen*, Vermeidung von Divisionen, Crelles Journal für die reine und angewandte Mathematik **264**(1973), 184–202.

[18] *V. Strassen*, The complexity of continued fraction, SIAM J. Comp. **12/1**(1983), 1–27.

[19] *V. Strassen*, Relative bilinear complexity and matrix multiplication, J. für die reine und angewandte Mathematik **375/376**(1987), 406–443.

[20] *I. Wegener*, The complexity of Boolean functions, Wiley–Teubner, 1987.