Theoretical Computer Science 22 (1983) 317–330
North-Holland Publishing Company

THE COMPLEXITY OF PARTIAL DERIVATIVES

Walter BAUR and Volker STRASSEN
Seminar für Angewandte Mathematik, Universität Zürich, CH-8032 Zürich, Switzerland

Communicated by A. Schönhage
Received January 1982

Abstract. Let L denote the nonscalar complexity in $k(x_1, \ldots, x_n)$. We prove $L(f, \partial f/\partial x_1, \ldots, \partial f/\partial x_n) \leq 3L(f)$. Using this we determine the complexity of single power sums, single elementary symmetric functions, the resultant and the discriminant as root functions, up to order of magnitude. Also we linearly reduce matrix inversion to computing the determinant.

1. Introduction

Let k be an infinite field, $x_1, \ldots, x_n$ be indeterminates over k. Given $f_1, \ldots, f_q \in k(\mathbf{x})$, let $L(f_1, \ldots, f_q)$ be the minimal number of nonscalar multiplications/divisions sufficient to compute $f_1, \ldots, f_q$ from $x_1, \ldots, x_n$ allowing additions/subtractions and multiplications by arbitrary scalars from k for free. $L(f_1, \ldots, f_q)$ is called the complexity of $f_1, \ldots, f_q$. (For details see e.g. Borodin and Munro [1], Strassen [6])

One way to obtain lower bounds for the complexity of a set $f_1, \ldots, f_q$ of quolynomials (=rational functions) is by the degree method (Strassen [7]). Unfortunately in the case of single quolynomials one gets only trivial results. An interesting recent paper of Schnorr [4] deals with this problem and extends the method to yield nontrivial lower bounds for certain single functions. In the present paper we reduce the complexity of a single quolynomial to that of several quolynomials by means of the following simple but surprising inequality

$$L\left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \leq 3L(f), \tag{1}$$

proved in a completely elementary way. Combining (1) with the degree bound in its original form we obtain rather sharp complexity bounds, such as

$$L\left(\sum_{i=1}^{n} x_i^m\right) \succcurlyeq n \log m,$$

$$L(\sigma_q) \succcurlyeq n \log \min(q, n-q),$$

0304-3975/83/0000–0000/$03.00

where σ_q is the qth elementary symmetric function in n variables,

$$L\left(\prod_{i\neq j}(x_i - x_j)\right) \asymp n \log n.$$

Here $\asymp$ means equality of order of magnitude.

Of course (1) is useful not only for proving lower bounds. It easily implies, e.g. that computing the inverse of a matrix is not much harder than computing its determinant. In this connection we remark that inequalities similar to (1) hold for other cost measures (e.g. when counting all operations).

Throughout the paper log means $\log_2$. We apply Bezout's theorem in the form of Bezout's inequality for affine space (see Heintz [2], Schnorr [4]).

2. Main result

Theorem 1. *Let $f \in k(\boldsymbol{x})$. Then*

$$L\left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \leqslant 3L(f).$$

For the proof we need the following

Lemma. *Let K be a field, $y_i \in K$ nonzero $(1 \leqslant i \leqslant s)$, $\alpha_{ij} \in K$ $(1 \leqslant j < i \leqslant s)$ and $z_1, \ldots, z_s$ indeterminates over K. Define $h_1, \ldots, h_s$ by*

$$h_i = y_i\left(\sum_{j=1}^{i-1} \alpha_{ij} h_j + z_i\right), \quad i \leqslant s. \tag{2}$$

Write

$$h_s = \sum_{\sigma=1}^{s} d_\sigma z_\sigma$$

with $d_\sigma \in K$. Then

$$\begin{aligned} d_s &= y_s, \\ d_j &= \left(\sum_{\sigma=j+1}^{s} d_\sigma \alpha_{\sigma j}\right) y_j \quad \text{for } j < s. \end{aligned} \tag{3}$$

Proof. Let $h_i = \sum_{\sigma=1}^{s} d_{i\sigma} z_\sigma$. Then from (2)

$$d_{i\sigma} = y_i\left(\sum_{j=1}^{i-1} \alpha_{ij} d_{j\sigma} + \delta_{i\sigma}\right). \tag{4}$$

We introduce $(s \times s)$-matrices

$$D = (d_{i\sigma}) = \begin{pmatrix} d_{11} & & 0 \\ & d_{22} & \\ & & \ddots \\ d_{s1} & \cdots & d_{ss} \end{pmatrix},$$

$$A = \begin{pmatrix} 0 & & & 0 \\ \alpha_{21} & 0 & & \\ \vdots & & \ddots & \\ \alpha_{s1} & \cdots & \alpha_{s\,s-1} & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 & & 0 \\ & \ddots & \\ 0 & & y_s \end{pmatrix}.$$

Then (4) is equivalent to

$$D = Y(AD + 1) \tag{5}$$

Since $d_{ii} = y_i \neq 0$, D and Y are invertible Multiplying (5) from the left by Y^{-1} and from the right by D^{-1} we get

$$Y^{-1} = A + D^{-1}.$$

Multiplication from the left by D and from the right by Y yields

$$D = (DA + 1)Y$$

This means

$$d_{ij} = \Big(\sum_{\sigma=j+1}^{s} d_{i\sigma}\alpha_{\sigma j} + \delta_{ij} \Big) y_j.$$

For $i = s$ this is (3). □

Proof of Theorem 1. Let $L(f) \leq r$ Then there is a sequence $g_1, \ldots, g_r \in k(\boldsymbol{x})$ such that for all $i \leq r$ we have $g_i = u_i * / v_i$, where

$$u_i = \sum_{j=1}^{i-1} \beta_{ij} g_j + p_i, \qquad v_i = \sum_{j=1}^{i-1} \gamma_{ij} g_j + q_i \tag{6}$$

for some $\beta_{ij}, \gamma_{ij} \in k$, $p_i, q_i \in k + \sum_{\nu=1}^{n} k x_\nu$, and such that

$$f = \sum_{j=1}^{r} \alpha_j g_j + m$$

for some $\alpha_j \in k$, $m \in k + \sum_{\nu=1}^{n} k x_\nu$ In addition we may assume that all u_i, v_i are nonzero

The proof now proceeds as follows: It will first be shown that any partial derivative $\partial f / \partial x_\nu$ $(\nu \leq n)$ is of the form

$$\frac{\partial f}{\partial x_\nu} = h_s(\zeta_{\nu 1}, \ldots, \zeta_{\nu s}) \tag{7}$$

where $h_s(z_1, \ldots, z_s)$ is defined as in the lemma (relative to suitable parameters s, α_{ij}, y_i) and the $\zeta_{\nu\sigma}$ are elements from the ground field k. Since scalar multiplications are free it will then be sufficient to show that the coefficients d_σ of h_s can be computed from $g_1, \ldots, g_r$ with at most $2r$ multiplications/divisions. This will be done according to the recursion (3).

Now fix $\nu \leqslant n$. Using (6) and Leibniz's formula we obtain for all $i \leqslant r$

$$\frac{\partial g_i}{\partial x_\nu} = v_i\left(\sum_{j=1}^{i-1} \beta_{ij}\frac{\partial g_j}{\partial x_\nu} + \frac{\partial p_i}{\partial x_\nu}\right) + u_i\left(\sum_{j=1}^{i-1} \gamma_{ij}\frac{\partial g_j}{\partial x_\nu} + \frac{\partial q_i}{\partial x_\nu}\right) \tag{8}$$

if $g_i = u_i * v_i$, and

$$\frac{\partial g_i}{\partial x_\nu} = \frac{1}{v_i}\left[\left(\sum_{j=1}^{i-1} \beta_{ij}\frac{\partial g_j}{\partial x_\nu} + \frac{\partial p_i}{\partial x_\nu}\right) + \left(-\frac{u_i}{v_i}\right)\left(\sum_{j=1}^{i-1} \gamma_{ij}\frac{\partial g_j}{\partial x_\nu} + \frac{\partial q_i}{\partial x_\nu}\right)\right] \tag{9}$$

if $g_i = u_i/v_i$. Moreover

$$\frac{\partial f}{\partial x_\nu} = \sum_{j=1}^{r} \alpha_j \frac{\partial g_j}{\partial x_\nu} + \frac{\partial m}{\partial x_\nu} \tag{10}$$

In order to get into the situation of the lemma put $s = 3r+1$, and define α_{tl} for $1 \leqslant l < t \leqslant s$ by

$$\alpha_{3i-2\,l} = \begin{cases} \beta_{ij} & \text{if } l = 3j, \\ 0 & \text{if } 3 \nmid l, \end{cases}$$

$$\alpha_{3i-1\,l} = \begin{cases} \gamma_{ij} & \text{if } l = 3j, \\ 0 & \text{if } 3 \nmid l, \end{cases} \tag{11}$$

$$\alpha_{3i\,l} = \begin{cases} 1 & \text{if } l = 3i-2, \\ 1 & \text{if } l = 3i-1,\ g_i = u_i * v_i, \\ -1 & \text{if } l = 3i-1,\ g_i = u_i/v_i, \\ 0 & \text{otherwise,} \end{cases}$$

for $i \leqslant r$, and

$$\alpha_{3r+1\,l} = \begin{cases} \alpha_j & \text{if } l = 3j, \\ 0 & \text{if } 3 \nmid l \end{cases}$$

Furthermore, put $y_s = 1$ and for $i \leqslant r$ put

$$y_{3i-2} = v_i, \qquad y_{3i-1} = u_i, \qquad y_{3i} = 1 \tag{12}$$

in case $g_i = u_i * v_i$, and

$$y_{3i-2} = 1, \qquad y_{3i-1} = \frac{u_i}{v_i}, \qquad y_{3i} = \frac{1}{v_i} \tag{13}$$

in case $g_i = u_i/v_i$. Note that all $y_i \neq 0$.

Finally put

$$\zeta_{3i-2}=\frac{\partial p_i}{\partial x_\nu},\quad \zeta_{3i-1}=\frac{\partial q_i}{\partial x_\nu},\quad \zeta_{3i}=0 \qquad (i\leqslant r), \tag{14}$$

and

$$\zeta_s=\frac{\partial m}{\partial x}.$$

Let $h_1, \ldots, h_s$ be defined as in the lemma.

Claim. *For all $i\leqslant r$ we have*

$$\frac{\partial g_i}{\partial x_\nu}=h_{3i}(\zeta_1,\ldots,\zeta_s)=h_{3i}(\zeta) \quad \textit{and} \quad \frac{\partial f}{\partial x_\nu}=h_s(\zeta)$$

The first assertion is proved by induction on $i\leqslant r$. We treat the case $g_i=u_i/v_i$, leaving the case $g_i=u_i * v_i$ and the second assertion to the reader.

$$\frac{\partial g_i}{\partial x_\nu}=\frac{1}{v_i}\left[\left(\sum_{j=1}^{i-1}\beta_{ij}\frac{\partial g_j}{\partial x_\nu}+\frac{\partial p_i}{\partial x_\nu}\right)+\left(-\frac{u_i}{v_i}\right)\left(\sum_{j=1}^{i-1}\gamma_{ij}\frac{\partial g_j}{\partial x_\nu}+\frac{\partial q_i}{\partial x_\nu}\right)\right] \quad \text{(by (9))}$$

$$=\frac{1}{v_i}\left[y_{3i-2}\left(\sum_{l=1}^{3i-3}\alpha_{3i-2,l}h_l(\zeta)+\zeta_{3i-2}\right)\right.$$

$$\left.+y_{3i-1}\left(\sum_{l=1}^{3i-2}\alpha_{3i-1,l}h_l(\zeta)+\zeta_{3i-1}\right)\right]$$

(by induction hypothesis, (11), (13), (14))

$$=\frac{1}{v_i}[h_{3i-2}(\zeta)+h_{3i-1}(\zeta)] \qquad \text{(by (2))}$$

$$=y_{3i}\left[\sum_{l=1}^{3i-1}\alpha_{3i,l}h_l(\zeta)+\zeta_{3i}\right] \qquad \text{(by (11), (13), (14))}$$

$$=h_{3i}(\zeta) \qquad \text{(by (2))}$$

It follows from our claim that any partial derivative $\partial f/\partial x_\nu$ $(\nu\leqslant n)$ is of the desired form (7). Therefore

$$L\left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\,\middle|\,g_1,\ldots,g_r\right)\leqslant L(d_1,\ldots,d_s\,|\,g_1,\ldots,g_r) \tag{15}$$

where $d_1,\ldots,d_s$ are the coefficients of h_s. By (3) each d_j $(j<s)$ is obtained from $d_{j+1},\ldots,d_s$ using just one (nonscalar) multiplication, one factor being y_j. By (12) and (13) multiplication by y_j means either multiplication or division by some element from $k+\sum_{\nu=1}^{n}kx_\nu+\sum_{i=1}^{r}kg_i$. Furthermore, at least r of these multiplications are

multiplications by 1, and also $y_s = 1$. Hence

$$L(d_1, \ldots, d_s \mid g_1, \ldots, g_r) \leqslant (s-1) - r = 2r.$$

This and (15) imply the theorem. □

To a reader who would like to improve the constant 3 in Theorem 1 we suggest to look at the example

$$f = \sum_{i=1}^{n} \frac{x_i}{y_i} \in k(x_1, \ldots, y_n)$$

3. Applications

Corollary 1. *Assume* char $k \nmid m$. *Then*

$$\tfrac{1}{3} n \log(m-1) \leqslant L\left(\sum_{i=1}^{n} x_i^m\right) \leqslant nl(m)$$

where $l(m)$ *is the length of a shortest addition chain for* m (*see* Knuth [3]; $l(m) \leqslant 2 \log m$ *always*, $l(m) \sim \log m$).

Proof. It suffices to show that the left inequality. The theorem yields

$$L\left(\sum_{i=1}^{n} x_i^m\right) \geqslant \tfrac{1}{3} L(x_1^{m-1}, \ldots, x_n^{m-1})$$

Now we apply the degree bound (Strassen [7]). W.l.o.g. k algebraically closed. In case char $k \nmid m-1$

$$\begin{aligned} &\deg \operatorname{graph}(x_1^{m-1}, \ldots, x_n^{m-1}) \\ &\quad \geqslant \#\{(\xi_1, \ldots, \xi_n) \in k^n \mid \xi_1^{m-1} = \cdots = \xi_n^{m-1} = 1\} \\ &\quad = (m-1)^n, \end{aligned}$$

hence $L(x_1^{m-1}, \ldots, x_n^{m-1}) \geqslant n \log(m-1)$. In case char $k \mid m-1$

$$\begin{aligned} &\deg \operatorname{graph}(x_1^{m-1} + x_1, \ldots, x_n^{m-1} + x_n) \\ &\quad \geqslant \#\{(\xi_1, \ldots, \xi_n) \mid \xi_1^{m-1} + \xi_1 = \cdots = \xi_n^{m-1} + \xi_n = 0\} \\ &\quad = (m-1)^n, \end{aligned}$$

hence $L(x_1^{m-1}, \ldots, x_n^{m-1}) = L(x_1^{m-1} + x_1, \ldots, x_n^{m-1} + x_1) \geqslant n \log(m-1)$. □

Corollary 2. *Let* $\sigma_1, \ldots, \sigma_n$ *be the elementary symmetric functions in* $x_1, \ldots, x_n$, $q \leqslant \frac{1}{2}n$. *Then*

$$\tfrac{1}{3}(n-q+1)\log(q-1) \leqslant L(\sigma_q) \leqslant n \log q + 2n, \tag{16}$$

and

$$|L(\sigma_{n-q}) - L(\sigma_q)| \leqslant 2n \tag{17}$$

Proof. Left inequality of (16): W.l.o.g k algebraically closed and of infinite degree of transcendency over its prime field. Denote by $\sigma_q^{(m)}$ the qth elementary symmetric function in m variables. Obviously

$$\frac{\partial \sigma_q^{(n)}}{\partial x_i} = \sigma_{q-1}^{(n-1)}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), \tag{18}$$

and

$$\sigma_j^{(n-1)}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$
$$= \sigma_j^{(n)}(x_1, \ldots, x_n) - x_i \sigma_{j-1}^{(n-1)}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

The last recursion yields

$$\sigma_{q-1}^{(n-1)}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$
$$= \sigma_{q-1}^{(n)} - x_i \sigma_{q-2}^{(n)} + x_i^2 \sigma_{q-3}^{(n)} - \cdots + (-1)^{q-1} x_i^{q-1}$$

This together with (18) gives

$$\frac{\partial \sigma_q}{\partial x_i} = \sigma_{q-1} - x_i \sigma_{q-2} + \cdots + (-1)^{q-1} x_i^{q-1} \tag{19}$$

The theorem tells us

$$L(\sigma_q) \geqslant \tfrac{1}{3} L\left(\frac{\partial \sigma_q}{\partial x_1}, \ldots, \frac{\partial \sigma_q}{\partial x_n}\right). \tag{20}$$

Again we apply the degree bound to the right-hand side of this inequality, getting

$$L\left(\frac{\partial \sigma_q}{\partial x_1}, \ldots, \frac{\partial \sigma_q}{\partial x_n}\right) \geqslant \log \deg W, \tag{21}$$

where $W \subseteq k^{2n}$ is the graph of the polynomial map given by the equations

$$y_1 = \frac{\partial \sigma_q}{\partial x_1}(x_1, \ldots, x_n)$$
$$\vdots$$
$$y_n = \frac{\partial \sigma_q}{\partial x_n}(x_1, \ldots, x_n)$$

Choose $\lambda_1, \ldots, \lambda_{q-1}, \mu_q, \ldots, \mu_n \in k$ algebraically independent and intersect W with the hypersurfaces

$$\begin{aligned} &\sigma_1(x_1, \ldots, x_n) = \lambda_1 \\ &\vdots \\ &\sigma_{q-1}(x_1, \ldots, x_n) = \lambda_{q-1} \end{aligned}$$

and with the hyperplanes

$$y_q = \mu_q, \ldots, y_n = \mu_n.$$

For the intersection W_0 we obtain by Bezout's inequality

$$\deg W_0 \leq \deg W \cdot (q-1)!$$

By (20) and (21) it suffices to show that W_0 is finite, and

$$\# W_0 = (q-1)^{n-q+1}(q-1)! \tag{22}$$

For $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in k^{2n}$ we have $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in W_0$ if and only if (23), (24) and (25), where

$$\eta_i = \frac{\partial \sigma_q}{\partial x_i}(\boldsymbol{\xi}), \quad 1 \leq i \leq n, \tag{23}$$

$$\sigma_i(\boldsymbol{\xi}) = \lambda_i, \quad 1 \leq i < q, \tag{24}$$

$$\xi_i^{q-1} - \lambda_1 \xi_i^{q-2} + \cdots + (-1)^{q-1}\lambda_{q-1} = (-1)^{q-1}\mu_i, \quad q \leq i \leq n, \tag{25}$$

the last group of equations coming from (19). Since $\lambda_1, \ldots, \lambda_{q-1}, \mu_q, \ldots, \mu_n$ are algebraically independent there are precisely $(q-1)^{n-q+1}$ vectors $(\xi_q, \ldots, \xi_n)$ satisfying (25). Hence it suffices to show that any such vector has precisely $(q-1)!$ extensions to a vector $(\xi_1, \ldots, \xi_n)$ satisfying (24).

Now fix $(\xi_q, \ldots, \xi_n)$ satisfying (25). Since obviously

$$\sum_{j=0}^{n} (-1)^j \sigma_j(\boldsymbol{\xi}) \xi_i^{n-j} = 0,$$

(24) and (25) imply

$$\sum_{j=q}^{n} (-1)^j \sigma_j(\boldsymbol{\xi}) \xi_i^{n-j} = (-1)^q \mu_i \xi_i^{n-q+1}, \quad q \leq i \leq n. \tag{26}$$

Since the components of $(\xi_q, \ldots, \xi_n)$ are pairwise different, the system of linear equations

$$\sum_{j=q}^{n} z_j (-1)^j \xi_i^{n-j} = (-1)^q \mu_i \xi_i^{n-q+1}, \quad q \leq i \leq n$$

has a unique solution $\zeta_q, \ldots, \zeta_n$. Therefore, using (26), an extension $(\xi_1, \ldots, \xi_n)$ of $(\xi_q, \ldots, \xi_n)$ satisfies (24) if and only if

$$\begin{aligned} &\sigma_i(\boldsymbol{\xi}) = \lambda_i \quad 1 \leq i < q, \\ &\sigma_j(\boldsymbol{\xi}) = \zeta_j \quad q \leq j \leq n. \end{aligned} \tag{27}$$

Introducing

$$f(t) = t^n - \lambda_1 t^{n-1} + \cdots + (-1)^{q-1}\lambda_{q-1}t^{n-q+1} + (-1)^q \zeta_q t^{n-q} + \cdots + (-1)^n \zeta_n,$$

(27) is equivalent to

$$f(t) = \prod_{i=1}^{n} (t - \xi_i). \tag{28}$$

Since by the definition of the ξ_i and by (25), $\xi_q, \ldots, \xi_n$ are roots of $f(t)$ there is at least one solution $(\xi_1, \ldots, \xi_n)$ (extending $(\xi_q, \ldots, \xi_n)$) of equation (28). Choosing any such solution we get all solutions by permuting the first $q-1$ components $\xi_1, \ldots, \xi_{q-1}$. So it remains to show that $\xi_1, \ldots, \xi_{q-1}$ are pairwise different. But in fact $\xi_1, \ldots, \xi_n$ are algebraically independent since by (24) and (25) $\lambda_1, \ldots, \lambda_{q-1}, \mu_q, \ldots, \mu_n$ are rational functions of them. Thus we have shown

$$L(\sigma_q) \geqslant \tfrac{1}{3}(n-q+1)\log(q-1).$$

Right inequality of (16). Let $m = \lfloor n/q \rfloor$ and $p = n - mq$. For any $i = 0, \ldots, m-1$ compute all the elementary symmetric functions $\sigma_{i1}, \ldots, \sigma_{iq}$ in $x_{iq+1}, \ldots, x_{(i+1)q}$, and all elementary symmetric functions $\sigma_{m,1}, \ldots, \sigma_{mp}$ in $x_{mq+1}, \ldots, x_n$. This can be done with cost

$$mq \log q + p \log p \leqslant n \log q$$

(see e.g. Strassen [7, p. 243]). We introduce the polynomials

$$Q_i = 1 - \sigma_{i1}t + \cdots + (-1)^q \sigma_{iq} t^q,$$

$$Q_m = 1 - \sigma_{m,1}t + \cdots + (-1)^p \sigma_{mp} t^p.$$

Then

$$\prod_{i=0}^{m} Q_i \equiv 1 - \sigma_1 t + \cdots + (-1)^q \sigma_q t^q \pmod{t^{q+1}}.$$

Since symbolic multiplication mod t^{q+1} of two polynomials with constant term 1 can be done with $2q$ nonscalar operations, we can compute $\sigma_1, \ldots, \sigma_q$ from the σ_{ij} in time $2qm \leqslant 2n$. Thus

$$L(\sigma_q) \leqslant L(\sigma_1, \ldots, \sigma_q) \leqslant n \log q + 2n.$$

(17) follows from the equation

$$\sigma_{n-q} = \sigma_q\left(\frac{1}{x_1}, \ldots, \frac{1}{x_n}\right) x_1 \cdots x_n,$$

valid for $1 \leqslant q \leqslant n$. □

An interesting consequence of Corollary 2 is the following: There is a polynomial

$$f(x_1, \ldots, x_{2n}, t) = \sum_{i=0}^{2n} f_i(x_1, \ldots, x_{2n}) t^i$$

such that

$$L(f) = 2n - 1, \qquad L(f_n) \geq \tfrac{1}{3} n \log(n-1).$$

(Take $f = (t - x_1) \cdots (t - x_{2n})$. Compare this also with Valiant [9;§4].)

Corollary 3 (resultant). *Let $x_1, \ldots, x_n, y_1, \ldots, y_n$ be indeterminates over k. Then*

$$\tfrac{1}{3} n \log n \leq L\left(\prod_{i,j} (x_i - y_i)\right) < n(9 \log n + 1)$$

Proof. *Left inequality*: It suffices to show that

$$L\left(\prod_{i,j} (x_i - \eta_j)\right) \geq \tfrac{1}{3} n \log n,$$

where $\eta_1, \ldots, \eta_n \in k$ are algebraically independent over the prime field. (Just adjoin the y_j to the ground field k.) Put

$$f = \prod_{i,j} (x_i - \eta_j), \qquad g = \prod_j (t - \eta_j),$$

where t is a new indeterminate. Then

$$f = \prod_{i=1}^{n} g(x_i), \tag{29}$$

$$\frac{\partial f}{\partial x_i} = f \frac{g'(x_i)}{g(x_i)}. \tag{30}$$

The theorem and the degree bound yield

$$L(f) \geq \tfrac{1}{3} L\left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \geq \tfrac{1}{3} \log \deg \text{graph } \phi,$$

where ϕ is the polynomial map defined by $f, \partial f/\partial x_1, \ldots, \partial f/\partial x_n$. We intersect graph ϕ, which lives in k^{2n+1} with coordinate variables $x_1, \ldots, x_n, z_0, \ldots, z_n$ (say), with the equations

$$z_0 = z_1, \ldots, z_0 = z_n$$

By (29) and (30) the intersection is

$$V = \{(\xi, \phi(\xi)) \mid \text{either } f(\xi) = 0 \text{ or } \forall i \; g(\xi_i) - g'(\xi_i) = 0\}$$

By Bezout's inequality,

$$\deg \text{graph } \phi \geq \deg V$$

Now g and therefore $g-g'$ have algebraically independent coefficients (except for the highest which is 1). Thus $g-g'$ has n distinct roots, i.e

$$\#\{(\xi, \phi(\xi)) \mid \forall i\, g(\xi_i) - g'(\xi_i) = 0\} = n^n.$$

Now (29) implies that the above n^n points satisfy $f(\xi) \neq 0$. This means that V contains n^n isolated points, so $\deg V \geq n^n$

Right inequality: First compute the elementary symmetric functions in $y_1, \ldots, y_n$ in time $n \log n$, next evaluate the polynomial $t^n - \sigma_1(\mathbf{y}) t^{n-1} + \cdots + (-1)^n \sigma_n(\mathbf{y})$ at the points $x_1, \ldots, x_n$ in time $8n \log n$ and finally multiply the values. (Compare Borodin and Munro [1]). □

We do not know whether our method allows to prove a nonlinear lower bound for the complexity of the resultant of two polynomials as a function of their coefficients.

Corollary 4 (discriminant) *Let $x_1, \ldots, x_n$ be indeterminates over k. Then*

$$\tfrac{1}{6} n \log n - \tfrac{2}{3} n < L\left(\prod_{i \neq j} (x_i - x_j)\right) < n(9 \log n + 1).$$

Proof. *Upper bound* (suggested by J. Stoss): First compute the coefficients of $A(t) = \Pi_{i=1}^n (t - x_i)$ with cost $n \log n$, then the coefficients of $\mathrm{d}A/\mathrm{d}t = \sum_{j=1}^n \Pi_{i \neq j} (t - x_i)$ without additional cost. Now evaluate $\mathrm{d}A/\mathrm{d}t$ at $x_1, \ldots, x_n$ and multiply the values. This can be done with cost $8n \log n + n - 1$ (actually $7n \log n + n - 1$ is sufficient, using byproducts of step 1).

Lower bound: Let $f = \Pi_{i \neq j} (x_i - x_j)$, $p = \lceil n/2 \rceil$, $q = \lfloor n/2 \rfloor$. For clarity replace $x_1, \ldots, x_n$ by $x_1, \ldots, x_p, y_1, \ldots, y_q$. Then

$$f = g(\mathbf{x}) \cdot \prod_{i\,j} (x_j - y_j)^2 \cdot h(\mathbf{y}),$$

and therefore

$$\frac{\partial f}{\partial y_j} = f \cdot \left(- \sum_{l=1}^{p} \frac{2}{x_l - y_j} + \frac{\partial h}{\partial y_j}(\mathbf{y}) \cdot \frac{1}{h(\mathbf{y})} \right).$$

The theorem yields

$$L\left(f, \frac{\partial f}{\partial y_1}, \ldots, \frac{\partial f}{\partial y_q}\right) \leq 3L(f).$$

Adjoining the y_j to the ground field and calling them η_j we get

$$L\left(f(\mathbf{x}, \boldsymbol{\eta}), f(\mathbf{x}, \boldsymbol{\eta}) \sum_{l=1}^{p} \frac{1}{x_l - \eta_1}, \ldots, f(\mathbf{x}, \boldsymbol{\eta}) \sum_{l=1}^{p} \frac{1}{x_l - \eta_q}\right) \leq 3L(f).$$

Dividing $f(x, \eta)$ by the other q terms under the bracket we obtain

$$L\left(\frac{\prod_{i=1}^{p}(x_i-\eta_1)}{\sum_{l=1}^{p}\prod_{i\neq l}(x_i-\eta_1)}, \ldots, \frac{\prod_{i=1}^{p}(x_i-\eta_q)}{\sum_{l=1}^{p}\prod_{i\neq l}(x_i-\eta_q)}\right) \leq 3L(f)+q.$$

If ϕ is the rational map defined by the above q (reduced) quolynomials, the degree bound yields

$$3L(f)+q \geq \log \deg \text{graph } \phi.$$

Now

$$\begin{aligned}\deg \text{graph } \phi &\geq \#\text{ of components of } \phi^{-1}(0)\\ &= \#\text{ of components of } \{\xi \in k^p \mid \forall j \exists i \xi_i = \eta_j\}\\ &= p(p-1) \cdots (p-q+1) = p!,\end{aligned}$$

because the η_j are pairwise distinct. Therefore

$$\begin{aligned}3L(f) &\geq \log p! - q \geq \frac{n}{2}\log\frac{n}{2e} - \frac{n}{2}\\ &> \frac{n}{2}\log n - 2n. \quad \square\end{aligned}$$

Corollary 4 implies that the complexity of the Vandermonde determinant $\Pi_{i<j}(x_i - x_j)$ is at least $\frac{1}{6}n \log n - n$ (Its square is the discriminant.)

Corollary 5. *Let a_{ij} $(1 \leq i, j \leq n)$ be indeterminates over k, $(b_{ij}) = (a_{ij})^{-1}$ as matrices. Then*

$$L(\{b_{ij} \mid 1 \leq i, j \leq n\}) \leq 3L(\det(a_{ij})) + n^2$$

Proof. By Cramer's rule

$$\frac{\partial}{\partial a_{ij}} \det(a) = b_{ji} \det(a)$$

Thus

$$\begin{aligned}L(\{b_{ij} \mid 1 \leq i, j \leq n\}) &\leq L\left(\{\det(a)\} \cup \left\{\frac{\partial}{\partial a_{ij}}\det(a) \mid 1 \leq i, j \leq n\right\}\right) + n^2\\ &\leq 3L(\det(a)) + n^2 \quad \square\end{aligned}$$

Corollary 5, in conjunction with e.g. Strassen [5, 8], shows that the determinant has roughly the same complexity as matrix multiplication or inversion. It would be interesting to have a similar result for solving a system of linear equations

Corollary 6. *Let* $f = \sum_{i,j,l} \tau_{ijl} x_i y_j z_l$ *be a trilinear form. Then*

$$L(f) \geq \tfrac{1}{6} \operatorname{rank}(\tau_{ijl}).$$

Proof. Differentiate with respect to the x_i and apply Korollar 3 and Lemma 6 of Strassen [8] □

4. An extension

Theorem 2. *Let* $f \in k(x_1, \ldots, x_n)$ *be computable from* $\{x_1, \ldots, x_n\} \cup k$ *using* A *additions/subtractions,* S *scalar multiplications (i.e. multiplications by elements from* k*) and* M *further multiplications/divisions. Then* $\{f, \partial f/\partial x_1, \ldots, \partial f/\partial x_n\}$ *can be computed from* $\{x_1, \ldots, x_n\} \cup k$ *using* $2A + M$ *additions/subtractions,* $2S$ *scalar multiplications and* $3M$ *further multiplications/divisions.*

In particular, let L_1 *denote the complexity when all operations have unit cost,* L_2 *the complexity when additions/subtractions are free but all multiplications/divisions (including scalar multiplications) count. Then*

$$L_1\left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \leq 4L_1(f),$$

$$L_2\left(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right) \leq 3L_2(f).$$

Sketch of proof of Theorem 2. We may assume that the given computation sequence is of the form $g_1, \ldots, g_{n+m+r}$, where $g_i = x_i$ for $1 \leq i \leq n$, $g_i \in k$ for $n < i \leq n+m$, and any g_i with $n+m < i \leq n+m+r$ is obtained by adding/subtracting or multiplying/dividing two previous g's or by multiplying a previous g by some element from k.

Proceed as in the proof of Theorem 1 with the following provisions:

(1) Each addition/subtraction or multiplication/division in the given computation yields three rows of the matrix (α_{ij}) as before. Each scalar multiplication as well as each of the $n+m$ initial steps give rise to only one row of (α_{ij}).

(2) All $\alpha_{ij} \in \{0, 1, -1\}$.

Since we may assume that in the original computation any intermediate result except the last one is being referred to, all columns of the matrix (α_{ij}) except the last one are nonzero

Now the lemma can be applied and (7) holds with $\zeta_{\nu i} = \delta_{\nu i}$. It is clear that the given computation together with the new one provided by the lemma use $2S$ scalar multiplications and $3M$ further multiplications/divisions. It remains to estimate the total number of additions/subtractions. If we content ourselves first with computing the d_i up to sign only, the number B of additions/subtractions used can be made equal to A plus the number of nonzero α_{ij} minus the number of nonzero

columns of the matrix (α_{ij}), i.e.

$$B = A + (4A + S + 4M) - (n + m + 3A + S + 3M - 1)$$
$$= 2A + M - (n + m - 1).$$

Now we can adjust the signs of $d_1, \ldots, d_{n+m}$ (we are interested only in $d_1, \ldots, d_n$) using at most $n + m - 1$ additional subtractions, since not all of $d_1, \ldots, d_{n+m}$ have the wrong sign. (To see this observe that the first i such that our procedure yields d_i with the correct sign cannot exceed $n + m$.) □

References

[1] A. Borodin and I. Munro, *The Computational Complexity of Algebraic and Numeric Problems* (American Elsevier, New York, 1975).

[2] J. Heintz, Definability bounds of first order theories of algebraically closed fields, extended abstract, in: L. Budach, Ed., *Fundamentals of Computation Theory FCT 79* (Akademie-Verlag, Berlin 1979) 160–166.

[3] D.E. Knuth, *The Art of Computer Programming Vol. II: Seminumerical Algorithms* (Addison-Wesley, Reading, MA, 1969).

[4] C.P. Schnorr, An extension of Strassen's degree bound, *SIAM J. Comput.* **10** (1981) 371–382.

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

[6] V. Strassen, Berechnung und Programm I, *Acta Informat.* **1** (1972) 320–335.

[7] V. Strassen, Die Berechnungskomplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten, *Numer. Math.* **20** (1973) 238–251.

[8] V. Strassen, Vermeidung von Divisionen, *Crelle J. Reine Angew. Math.* **264** (1973) 184–202.

[9] L.G. Valiant, Reducibility by algebraic projections, *Logic and Algorithmic*, an International Symposium held in honour of Ernst Specker, Genève, *L'Enseignement Math.* (1982).