J. Symbolic Computation (1995) **11**, 1–18

On the search of the maximum of a polynomial

ALEXEI YU. UTESHEV † AND TIMOFEI M. CHERKASOV ‡

† *Department of Applied Mathematics, St.-Petersburg State University,*

Bibliotechnaja pl.2, 198904, St.-Petersburg, RUSSIA

e-mail: irina@utesh.spb.su ; uteschew@risc.uni-linz.ac.at (till 31.07.95)

‡ *Research Institute for Symbolic Computation, Johannes Kepler University,*

RISC-Linz, Schloss Hagenberg, A-4232, AUSTRIA

e-mail: tcherk@risc.uni-linz.ac.at

(Received 1 September 1994)

For a real polynomial $f(X)$ of k variables the problem of finding $\max_{X\in\mathbb{R}^k} f(X)$ is investigated by reducing it to that of searching for the real roots of a polynomial in one variable: $\mathcal{F}(z) := \prod_{j=1}^{N}(f(X_j) - z)$,where X_j is a critical point of $f(X)$. The use of Hermite method of separation of real solutions of algebraic equation system permits one to construct together with $\mathcal{F}(z)$ also its Sturm series, and restore the coordinates of corresponding critical point.The problem of finding the max f in the set defined by the real polynomial inequality $G(X) \geq 0$ is also discussed.

1. Introduction.

The present paper deals with the problem which is a part of general polynomial optimization problem: finding the maximum of a polynomial function in the domain defined by polynomial inequality system. This nonlinear programming problem has been intensively treated by several authors during the last two decades, for survey and recent developments in this area we refer to Bank, Heintz, Krick, Mandel and Solernó (1992). We shall investigate here mainly the problem of finding

$$\max_{X\in\mathbb{R}^k} f(X) \quad \text{and} \quad \max_{X\in\mathbb{R}^k|G(X)\geq 0} f(X)$$

for polynomial functions f and G. The proposed approach is based on the Hermite method for the separation of solutions of systems of algebraic equations (Hermite (1912))

$$F_j(X) = 0, j = \overline{1,k} \tag{1.1}$$

i.e. , for finding the number of solutions of (1.1) which satisfy the inequality $G(X) \geq 0$. During the last years the method was renewed by several authors: Becker and Wörmann (1994); Pedersen, Roy and Szpirglas (1993); Uteshev and Shulyak (1992). Although it was sufficiently illuminated in those (and some other) papers, we decided to reconsider its backgrounds because of the following reasons.

0747–7171/90/000000 + 00 $03.00/0

Briefly, the idea of the method consists in finding the rank and the signature of the quadratic form in real $x_0, \ldots, x_{N-1}$:

$$\sum_{j=1}^{N} G(\Lambda_j)[x_0\Psi_0(\Lambda_j) + \ldots + x_{N-1}\Psi_{N-1}(\Lambda_j)]^2 \tag{1.2}$$

Here $\Lambda_1, \ldots, \Lambda_N$ are the solutions of (1.1), and the system of monomials $\{\Psi_j(\Lambda)\}_{j=0}^{N-1}$ is linearly independent on those solutions. Elements of the matrix of the form (1.2) are the symmetric polynomials of solutions. To evaluate them, the elimination procedure should be used. That is usually done with the help of (sub)resultant theory or by applying the Shape lemma from the Gröbner basis theory. We intend to make this procedure self-contained, i.e. to find the resultant via the symmetric functions of solutions.

Furthermore, the original Hermite's paper contains an interesting method for choosing the monomial system $\{\Psi_j(\Lambda)\}_{j=0}^{N-1}$. In general case, this can be done by considering only the leading forms in the expansions of polynomials $F_j(X)$. As far as we know, that fact has not been yet mentioned.

The problem of finding the $\max f(X)$ (if it $\neq +\infty$) can be treated in the ideology of Hermite method. If we consider the problem of finding the number of real solutions of the system

$$\partial f/\partial x_1 = 0, \ldots, \partial f/\partial x_k = 0$$

in the domain $f(X) > z$ (and $G(X) \geq 0$), we shall get the results mentioned in the abstract.

Notation:

1. For the matrix S : $r(S)$ denotes its rank, $\sigma(S)$ its signature, $n_+(S)$ its positive index, S_j its leading principal minor of the order j.
2. For the real polynomials $F(x)$ and $G(x)$: $\mathcal{D}(F)$ is the discriminant of $F(x)$, $\mathcal{R}(F, G)$ is the resultant of F and G, $\mathrm{nrr}\{F = 0 | G > 0\}$ is the number of distinct real roots of $F(x)$ which satisfy the inequality $G(x) > 0$.
3. For the real polynomials $F_1(x, y), F_2(x, y)$ and $G(x, y)$: $\mathcal{J}(F_1, F_2)$ (or $\mathcal{J}(x, y)$) is the Jacobian of F_1 and F_2, $\mathcal{X}(x)$ (or $\mathcal{Y}(y)$) is the resultant of F_1 and F_2 obtained on elimination of y (or x); $\mathrm{nrs}\{F_1 = 0, F_2 = 0 | G(x, y) > 0\}$ stands for the number of distinct real solutions of the system of equations satisfying the inequality $G(x, y) > 0$.

2. The case of one variable

We shall first present some preliminary results on the relative distribution of the roots of two polynomials in terms of the appropriate Hankel matrices (Krein, Naimark(1981); Uteshev, Shulyak(1992)). For the real polynomials

$$F(x) := A_0x^n + \cdots + A_n \quad \text{and} \quad G(x) := B_0x^m + \cdots + B_m$$

consider the following Laurent expansions

$$\frac{F'(x)}{F(x)} = \sum_{k=0}^{\infty} \frac{s_k}{x^{k+1}}$$

$$\frac{G(x)}{F(x)} = L(x) + \sum_{k=0}^{\infty} \frac{c_k}{x^{k+1}} \tag{2.1}$$

$$\frac{G(x)F'(x)}{F(x)} = L_1(x) + \sum_{k=0}^{\infty} \frac{h_k}{x^{k+1}} \tag{2.2}$$

We shall be interested mainly in the coefficients of the principal parts.To find them, it is convenient to construct first the expansion

$$\frac{1}{F(x)} = \sum_{k=n-1}^{\infty} \frac{d_k}{x^{k+1}} \tag{2.3}$$

with the coefficients determined by the recurrent formulae

$$d_k := \begin{cases} 1/A_0 \text{ if } \quad k = n-1 \\ \\ -A_1 d_{k-1}/A_0 \text{ if } \quad k = n \\ \\ -\dfrac{A_1 d_{k-1} + A_2 d_{k-2} + \cdots + A_{k-n+1} d_{n-1}}{A_0} \text{ if } \quad k > n \end{cases}$$

($A_k := 0$ if $k > n$), and then to multiply (2.3) by the corresponding polynomials. In this way, one can find the expressions for s_k, c_k and h_k which will be the rational functions with respect to the coefficients of F and G. On the other hand, those coefficients turn out to be symmetric polynomials of the roots $\lambda_1, \ldots, \lambda_n$ of $F(x)$. So, for example, s_k is known as the Newton sum:

$$s_k := \sum_{j=1}^{n} \lambda_j^k = \begin{cases} n \text{ if } \quad k = 0 \\ \\ -A_1/A_0 \text{ if } \quad k = 1 \\ \\ -\dfrac{A_1 s_{k-1} + A_2 s_{k-2} + \cdots + A_{k-1} s_1 + A_k k}{A_0} \text{ if } \quad k > 1 \end{cases}$$

($A_k := 0$ if $k > n$), and

$$h_k := \sum_{j=1}^{n} G(\lambda_j) \lambda_j^k = B_0 s_{k+m} + B_1 s_{k+m-1} + \cdots + B_m s_k$$

Consider the following Hankel matrices

$$S = [s_{j+k}]_{j,k=0}^{n-1}, C = [c_{j+k}]_{j,k=0}^{n-1}, H = [h_{j+k}]_{j,k=0}^{n-1} \tag{2.4}$$

THEOREM 2.1 (Jacobi). *The number of distinct roots of $F(x)$ equals $r(S)$, and the number of distinct real roots of $F(x)$ equals $\sigma(S)$.*

THEOREM 2.2 (Hermite, Sylvester). *If $\mathcal{R}(F,G) \neq 0$, then the number of distinct real roots of $F(x)$ satisfying the inequality $G(x) > 0$ equals*

$$n_+(H) - [r(S) - \sigma(S)]/2.$$

We shall denote that number by $\mathrm{nrr}\{F = 0 | G > 0\}$.

COROLLARY 2.1.*One has*

$$S_n := \det S = \prod_{1 \le j < k \le n} (\lambda_k - \lambda_j)^2 = \mathcal{D}(F)/A_0^{2n-2} \tag{2.5}$$

$$H_n := \det H = S_n \prod_{1 \le j \le n} G(\lambda_j) = \mathcal{R}(F,G) S_n / A_0^m. \tag{2.6}$$

$F(x)$ does not have multiple roots iff $S_n \neq 0$. Under this condition, $\mathcal{R}(F,G) \neq 0$ iff $H_n \neq 0$.

COROLLARY 2.2. *If all the leading principal minors S_j, H_k of the matrices S and H do not vanish, then*

$$\operatorname{nrr}\{F=0\} = P(1,S_1,\dots,S_n) - V(1,S_1,\dots,S_n)$$

$$\operatorname{nrr}\{F=0|G>0\} = P(1,H_1,\dots,H_n) - V(1,S_1,\dots,S_n),$$

where P [or V] is the number of permanences [or variations] of sign in the corresponding sequences; otherwise the rule of Frobenius (Gantmacher (1959)) *should be used.*

THEOREM 2.3 (Kronecker(1897)).
$\deg(\gcd(F,G)) = d$ *iff* $C_n = 0, \dots, C_{n-d+1} = 0, C_{n-d} \neq 0$. *In this case,* $\gcd(F,G)$ *equals the determinant obtained on replacing the last row in* C_{n-d}

$$\gcd(F,G) \equiv \begin{vmatrix} c_0 & c_1 & \dots & c_{n-d-1} \\ c_1 & c_2 & \dots & c_{n-d} \\ \vdots & \vdots & & \vdots \\ c_{n-d-2} & c_{n-d-1} & \dots & c_{2n-2d-3} \\ \sum_{j=n-d}^{n} c_{j-1}F_j & \sum_{j=n-d}^{n} c_jF_j & \dots & \sum_{j=n-d}^{n} c_{j+n-d-2}F_j \end{vmatrix}$$

where $F_k := A_0x^{n-k} + A_1x^{n-k-1} + \cdots + A_{n-k}$. If $n > m$, then polynomials $p(x)$ and $q(x)$ providing the linear representation of gcd

$$\gcd(F,G) \equiv p(x)F(x) + q(x)G(x)$$

may be expressed as the determinants obtained on replacing the last row of C_{n-d} by

$$[0, -p_0(x), -p_1(x), \dots, -p_{n-d-2}(x)] \text{ and } [1, x, \dots, x^{n-d-1}]$$

correspondingly. Here $p_k := c_0x^k + \dots + c_k$.

PROOF. We shall give here only the idea of the proof, since it will be used for the multivariable case. For the sake of siplicity, let $d = 0$, i.e. $gcd(F,G) \equiv const$. According with the statement of the theorem, we have to find the coefficients of polynomials

$$p(x) := \rho_0x^{n-2} + \rho_1x^{n-3} + \dots + \rho_{n-2}, q(x) := q_0x^{n-1} + q_1x^{n-2} + \dots + q_{n-1}$$

satisfying the equality $p(x)F(x) + q(x)G(x) \equiv a_0C_n$. On dividing the latter by $F(x)$ and expanding its both sides in series we get

$$p(x) + q(x)(\frac{c_0}{x} + \frac{c_1}{x} + \dots) \equiv a_0C_n(\frac{d_{n-1}}{x^n} + \frac{d_n}{x^{n+1}} + \dots).$$

Comparing coefficients of the equal powers of x gives a linear system with respect to $q_{n-1}, \dots, q_0$:

$$\begin{array}{ll} x^{-1} & : c_0q_{n-1} + c_1q_{n-2} + \dots + c_{n-1}q_0 = 0, \\ & \dots \\ x^{-n+1} & : c_{n-2}q_{n-1} + c_{n-1}q_{n-2} + \dots + c_{2n-3}q_0 = 0, \\ x^{-n} & : c_{n-1}q_{n-1} + c_nq_{n-2} + \dots + c_{2n-2}q_0 = C_n \end{array}$$

Solving this systems by Cramer's rule, we get the expression for $q_0, \dots, q_{n-1}$, which coincide with those mentioned in Theorem 2.3. Similary, one has

$$\begin{aligned} x^{n-2} &: \rho_0 + q_0 c_0 = 0, \\ x^{n-3} &: \rho_1 + q_0 c_1 + q_1 c_0 = 0, \\ &\dots \\ 1 &: \rho_{n-2} + q_0 c_{n-2} + q_1 c_{n-3} + \dots + q_{n-2} c_0 = 0 \end{aligned} \tag{2.7}$$

From the above equations, the coefficients $\rho_0, ..., \rho_{n-2}$ may be expressed in the determinantial form with the help of the representations of $q_0, ..., q_{n-1}$ obtained earlier. However, for practical calculations of $\rho_0, ..., \rho_{n-1}$ it would be easier to use the formulae (2.7) directly.

To conclude the proof, let us mention the fact that $p(x)$ equals the quotient on division of $q(x)g(x)$ by $f(x)$, multiplied by (-1). □

COROLLARY 2.3 *$A_0^{2k-(n-m)} C_k$ equals the kth subresultant of F and G, in particular case $A_0^{n+m} C_n = \mathcal{R}(F, G)$. By setting $d = 0$ in Theorem 2.3 one can also find the linear representation of $\mathcal{R}(F, G)$:*

$$\mathcal{R}(F, G) \equiv A_0^{n+m-1}(p(x)F(x) + q(x)G(x)) \tag{2.8}$$

provided that $m < n$. Otherwise, one should take into account the quotient $L(x)$ from division of $G(x)$ by $F(x)$ (the coefficients of $L(x)$ can be found also with the help of (2.3)).

COROLLARY 2.4 *For the case $d = 1$, the single common root of $F(x)$ and $G(x)$ can be represented as*

$$x = -\frac{A_1}{A_0} - \left| \begin{array}{cccc} c_0 & c_1 & \dots & c_{n-2} \\ c_1 & c_2 & \dots & c_{n-1} \\ \dots & & & \\ c_{n-3} & c_{n-2} & \dots & c_{2n-5} \\ c_{n-1} & c_n & \dots & c_{2n-3} \end{array} \right| / C_{n-1} \tag{2.9}$$

Moreover, since the expansion (2.2) is a particular case of (2.1), one may use in (2.9) also the corresponding minors of the matrix H provided that $S_n \neq 0$ (v. Corollary 2.1)

Resume. Theorems 2.1-2.3 permit one to solve the problem of establishing the relative distribution of the roots of two polynomials that is usually treated in terms of Euclid remainders or subresultants (Collins(1971); Gonzalez-Vega, Lombardi, Recio, Roy(1990)).

Consider now the problem of finding the maximum of a polynomial $f(x)$. On eliminating the trivial case $\max f(x) = +\infty$, we shall restrict ourselves to the polynomial of the even degree $n + 1$ with the negative leading coefficient:

$$f(x) := a_0 x^{n+1} + a_1 x^n + \ldots + a_n, \ a_0 < 0$$

PROBLEM 2.1. For every real z find

$$\mathrm{nrr}\{f'(x) = 0 | f(x) > z\}.$$

Let us use the Theorems 2.1-2.3. Take $F(x) := f'(x), G(x) := f(x) - z$ and construct the matrix H (2.4). Consider the sequence of its leading principal minors

$$H_1(z), \ldots, H_n(z) \tag{2.10}$$

Let us investigate the features of this sequence. First of all, the constant term of the polynomial $H_k(z)$ equals $(-1)^k S_k$, where S_k is the kth leading principal minor of the matrix S (2.4) constructed for $F(x) \equiv f'(x)$. So, by Theorem 2.1, one may find $\operatorname{nrr}\{f'(x)=0\}$.

ASSUMPTION 2.1. Let $S_n \neq 0$.

Thus, by Corollary 2.1, all the roots of $f'(x)$ are distinct. By (2.6), $H_n(z) = S_n \prod_{j=1}^n (f(\lambda_j) - z)$, and the polynomial

$$\mathcal{F}(z) := (-a_0)^n H_n(z)/S_n = (-a_0)^n \prod_{j=1}^{n} (f(\lambda_j) - z) \tag{2.11}$$

has the roots which coincide with the values of $f(x)$ at the roots $\lambda_1, \dots, \lambda_n$ of its derivative. By construction, the coefficients of $\mathcal{F}$ are polynomials of the coefficients of $f(x)$ with the constant term $(-1)^n \mathcal{D}(f)$.

EXAMPLE 2.1. If $f(x) = a_0x^3 + 3a_1x^2 + 3a_2x + a_3$, then

$$\mathcal{F}(z) = a_0^2 z^2 + (6a_0a_1a_2 - 4a_1^3 - 2a_0^2a_3)z - \mathcal{D}(f),$$

where $\mathcal{D}(f) = 4(a_1a_3 - a_2^2)(a_0a_2 - a_1^2) - (a_1a_2 - a_0a_3)^2$.

ASSUMPTION 2.2. Let $\mathcal{D}(\mathcal{F}(z)) \neq 0$.

Then all the critical values of $f(x)$ are distinct, and, by Corollary 2.2, we have the following result:

THEOREM 2.4. *Under the Assumptions* 2.1 *and* 2.2 *one has*

$$\operatorname{nrr}\{\mathcal{F}(z) = 0\} = \operatorname{nrr}\{f'(x) = 0\};$$

$$\operatorname{nrr}\{\mathcal{F}(z) = 0 | a < z < b\} = \operatorname{nrr}\{f'(x) = 0 | a < f(x) < b\}$$

$$= P(1, H_1(a), \dots, H_n(a)) - P(1, H_1(b), \dots, H_n(b)).$$

This theorem implies that the sequence (2.10) is the Sturm's series for $\mathcal{F}(z)$, by means of which the critical values of $f(x)$ may be localized. It is evident that $\max f(x)$ coincides with the greatest (real) root of $\mathcal{F}(z)$. Once it is evaluated, the corresponding value of x may be found also via the minors of the matrix $H(z)$. Indeed, by Corollary 2.4, x can be expressed by the formula (2.9), in which $A_1 := na_1, A_2 := (n+1)a_0, c_k := h_k(z)$, and $C_{n-1} := H_{n-1}(z)$.

It is evident that Assumption 2.2 is of the most importance for the proposed algorithm. It would be desirable to check it in terms of the matrices S and $H(z)$ (i.e. without an additional construction of the matrix S for the polynomial $\mathcal{F}(z)$). Our successes in determining the structure of $\mathcal{D}(\mathcal{F}(z))$ are restricted to the following

THEOREM 2.5. $\mathcal{D}(\mathcal{F}(z)) = k[\mathcal{D}(f')]^3 \Phi^2(A_0, \dots, A_n)$, *where k is a constant depending only on n, and Φ is a real homogeneous polynomial function of the degree* $3(n-1)(n-2)/2$ *with respect to the coefficients of $f'(x)$.*

PROOF. If $f(x) = a_0x^{n+1} + \dots + a_{n+1}$, $f'(x) = A_0x^n + \dots + A_n$ and $f'(\lambda_j) = f'(\lambda_k) = 0$, then

$$f(\lambda_k) - f(\lambda_j) = (\lambda_k - \lambda_j)^3 \Psi(\lambda_j, \lambda_k), \text{ where}$$

$$\Psi(\lambda_j,\lambda_k) := 1/2[a_{n-2} + a_{n-3}(2\lambda_j + 2\lambda_k)$$

$$+a_{n-4}(3\lambda_j^2 + 4\lambda_j\lambda_k + 3\lambda_k^2)$$

$$+a_{n-5}(4\lambda_j^3 + 6\lambda_j^2\lambda_k + 6\lambda_j\lambda_k^2 + 4\lambda_k^3) + \ldots]$$

(To understand the generative rule for the coefficients in parentheses put them in a triangle like Pascal's one...) Polynomial $\Psi(\lambda_j,\lambda_k)$ is a symmetric one with respect to λ_j,λ_k. Consequently, $\prod_{1\le j<k\le n}\Psi(\lambda_j,\lambda_k)$ is of the same type with respect to the roots of $f'(x)$ and thus can be expressed rationally in terms of $A_0,\ldots,A_n$. Hence, we have

$$\mathcal{D}(\mathcal{F}(z)) = (a_0^n)^{2n-2}\prod_{1\le j<k\le n}(f(\lambda_k)-f(\lambda_j))^2$$

$$= \frac{1}{(n+1)^{2n(n-1)}}\left(A_0^{2n-2}\prod_{1\le j<k\le n}(\lambda_k-\lambda_j)^2\right)^3\left(A_0^{(n-3)(n-1)}\prod_{1\le j<k\le n}\Psi(\lambda_j,\lambda_k)\right)^2$$

$$= k[\mathcal{D}(f'(x))]^3\Phi^2(A_0,\ldots,A_n)$$

Here Φ is a homogeneous polynomial of the degree $3(n-1)(n-2)/2$ and of the weight $(n-1)(n-2)/2$ with respect to $A_0,\ldots,A_n$; its coefficients may be made integers by an appropriate choice of the constant k :

for $n=2$, $\Phi\equiv 1$;

for $n=3$, $\Phi\equiv 27A_0^2A_3 + 2A_1^3 - 9A_0A_1A_2$;

for $n=4$, $\Phi\equiv A_0^6A_4^3 - 48A_0^5A_1A_3A_4^2 - 52A_0^5A_2^2A_4^2 + 36A_0^5A_2A_3^2A_4 + \ldots + 500A_2^6A_3^3$

(Φ contains 30 terms).$\square$

3. The case of two variables

As in the previous section, we shall begin with the sketch for the method of the separation of the solutions of a system of algebraic equations, which was originated in the works by Ch.Hermite (Hermite (1912); Becker and Wörmann (1994); Pedersen, Roy and Szpirglas (1993); Uteshev and Shulyak (1992)).

Let a system of polynomial equations be given

$$F_1(x,y)=0, F_2(x,y)=0 \tag{3.1}$$

where $\deg F_j := n_j$, and the real polynomial $G(x,y)$. The problem is to find the number of distinct real solutions of (3.1), that satisfy the inequality $G(x,y)>0$:

$$\mathrm{nrs}\{(3.1)|G(x,y)>0\}$$

Consider the leading forms in the expansions of F_1 and F_2 in decreasing powers of x and y

$$F_{j,n_j}(x,y) := A_{j,0}x^{n_j} + A_{j,1}x^{n_j-1}y + \ldots + A_{j,n_j}y^{n_j} \tag{3.2}$$

We shall suppose that $A_{j,0}\ne 0, A_{j,n_j}\ne 0$ for $j=1,2$.

Consider the resultants of F_1 and F_2 obtained on elimination of y and x respectively:

$$\mathcal{X}(x) := \mathcal{R}_y(F_1,F_2),\ \mathcal{Y}(y) := \mathcal{R}_x(F_1,F_2)$$

(also known as the *eliminants*). To find them, one may use Corollary 2.3. Thus, for example, to make $\mathcal{Y}(y)$ consider the fraction F_2/F_1 as a function of x and expand it in powers of x^{-1} :

$$\frac{F_2(x,y)}{F_1(x,y)} = L(x,y) + \sum_{k=0}^{\infty} \frac{c_k(y)}{x^{k+1}}$$

Then $\mathcal{Y}(y) \equiv A_{1,0}^{n_1+n_2} \det[c_{j+k}(y)]_{j,k=0}^{n_1-1}$. Moreover, by Corollary 2.3, one can also find its linear representation, i.e. , the polynomials $\mathcal{P}(x,y)$ and $\mathcal{Q}(x,y)$ such that

$$\mathcal{P}(x,y)F_1 + \mathcal{Q}(x,y)F_2 \equiv \mathcal{Y}(y) \tag{3.3}$$

$\mathcal{X}(x)$ can be constructed similarly together with polynomials $\mathcal{M}(x,y)$ and $\mathcal{N}(x,y)$ such that

$$\mathcal{M}(x,y)F_1 + \mathcal{N}(x,y)F_2 \equiv \mathcal{X}(x) \tag{3.4}$$

Generally, $\deg \mathcal{X} = \deg \mathcal{Y} = N$, where $N := n_1 n_2$ and

$$\mathcal{X}(x) = A_0 x^N + \textit{ lower order terms}, \ \mathcal{Y}(y) = (-1)^N A_0 y^N + \textit{lower order terms.}$$

Here

$$A_0 := \mathcal{R}(F_{1,n_1}(1,y), F_{2,n_2}(1,y)) \tag{3.5}$$

ASSUMPTION 3.1. Let $A_0 \neq 0$.

Under this assumption, the number of solutions of (3.1) (according with their multiplicities) equals N. If (α_j, β_j) is a solution of (3.1) then $\mathcal{X}(\alpha_j) = 0, \mathcal{Y}(\beta_j) = 0$ and vice versa: to every root of $\mathcal{X}$ corresponds at least one of the roots of $\mathcal{Y}$, such that this pair is a solution of (3.1).

Consider the function

$$\mathcal{V} := \mathcal{M}\mathcal{Q} - \mathcal{N}\mathcal{P}$$

and the expansion of the following fraction in the negative powers of x and y :

$$\frac{\mathcal{V}(x,y)}{\mathcal{X}(x)\mathcal{Y}(y)} = \sum_{j,k=0}^{\infty} \frac{d_{jk}}{x^{j+1}y^{k+1}} \tag{3.6}$$

(using the expansion of the type (2.3)). It could be proved (Uteshev and Shulyak (1992)) that

$$deg\mathcal{V} \leq 2N - n_1 - n_2, \quad deg_x\mathcal{V} \leq N-1, \quad deg_y\mathcal{V} \leq N-1 \tag{3.7}$$

$$d_{jk} = 0 \text{ for } j+k \leq n_1+n_2-3 \tag{3.8}$$

It turns out that the expansion (3.6) plays the role similar to that of the expansion (2.3). Indeed, let us multiply (3.6) first by the Jacobian $\mathcal{J}(x,y) := \partial F_1/\partial x \partial F_2/\partial y - \partial F_2/\partial x \partial F_1/\partial y$, then by $G(x,y)$, and, finally, by their product:

$$\frac{\mathcal{J}\mathcal{V}}{\mathcal{X}\mathcal{Y}} = L_0(x,y) + \sum_{j,k=0}^{\infty} \frac{s_{jk}}{x^{j+1}y^{k+1}} \ , \ \frac{G\mathcal{V}}{\mathcal{X}\mathcal{Y}} = L_1(x,y) + \sum_{j,k=0}^{\infty} \frac{c_{jk}}{x^{j+1}y^{k+1}} \tag{3.9}$$

$$\frac{\mathcal{J}G\mathcal{V}}{\mathcal{X}\mathcal{Y}} = L_2(x,y) + \sum_{j,k=0}^{\infty} \frac{h_{jk}}{x^{j+1}y^{k+1}}$$

Here L_j is of the form $[A(x,y)\mathcal{X}(x) + B(x,y)\mathcal{Y}(y)]/[\mathcal{X}\mathcal{Y}]$, where A and B are polynomials. We will be interested, however, only in the coefficients of the terms with negative powers of both variables.

THEOREM 3.1. (Jacobi) $s_{jk} = \sum_{p=1}^{N} \alpha_p^j \beta_p^k, h_{jk} = \sum_{p=1}^{N} G(\alpha_p, \beta_p)\alpha_p^j \beta_p^k$.
If $\mathcal{J}(\alpha_p, \beta_p) \neq 0$ *for all* $p = \overline{1, N}$ *(*i.e. *, all solutions of* (3.1) *are simple), then*

$$d_{jk} = \sum_{p=1}^{N} \frac{\alpha_p^j \beta_p^k}{\mathcal{J}(\alpha_p, \beta_p)} \tag{3.10}$$

and $= 0$ *for* $j + k \leq n_1 + n_2 - 3$ *(v.* (3.8)*).*

Thus, by Theorem 3.1 and by the construction of the expansions, s_{jk} and h_{jk} are the symmetric polynomials of solutions of (3.1), and so they can be expressed rationally in terms of the coefficients of F_1,F_2 and G.

To find now $\mathrm{nrs}\{(3.1)|G(x,y) > 0\}$ we shall consider two different approaches. The first one is based on the following

ASSUMPTION 3.2. Let $\mathcal{D}(\mathcal{X}(x)) \neq 0$.

Under this assumption, all the solutions (α_j, β_j) of (3.1) are distinct, they have distinct x–components, and

$$\mathrm{nrs}\{(3.1)\} = \mathrm{nrr}\{\mathcal{X}(x) = 0\}$$

where the latter number can be found by the Theorem 2.1.

THEOREM 3.2. (Brioschi) *Under the Assumptions 3.1 and 3.2, one has*

$$\mathrm{nrs}\{(3.1)|G(x,y) > 0\} = n_+(H) - (N - \mathrm{nrs}\{(3.1)\})/2, \tag{3.11}$$

where

$$H = [h_{j+k,0}]_{j,k=0}^{N-1} \tag{3.12}$$

COROLLARY 3.1. *One has:*

$$H_N := \det H = \prod_{1 \leq j < k \leq N} (\alpha_k - \alpha_j)^2 \prod_{1 \leq j \leq N} G(\alpha_j, \beta_j) = \frac{\mathcal{D}(\mathcal{X}(x))}{A_0^{2N-2}} \prod_{1 \leq j \leq N} G(\alpha_j, \beta_j).$$

In the original work by Hermite the idea of another approach can be also found (Uteshev and Cherkasov (1995)). In fact, it is possible to replace Assumption 3.2 by a simpler one, concerning only the forms (3.2). We shall sketch that approach only for a particular case when $n_1 = n_2 := n$.

ASSUMPTION 3.2$'$. Let all the subresultants $\mathcal{R}_1(A_0), ..., \mathcal{R}_{n-1}(A_0)$ of the resultant (3.5) be nonzero.

Construct the block-Hankel $N \times N$–matrices S, C and H. For example,

$$H := [H_{pl}]_{p,l=0}^{n-1}, \quad \text{where } H_{pl} := [h_{j+k,p+l}]_{j=\overline{0,2(n-p-1)};k=\overline{0,2(n-l-1)}} \tag{3.13}$$

S and C are constructed similarly from the coefficients of the expansions (3.9).

THEOREM 3.3. *Under Assumptions* 3.1 *and* 3.2$'$*, one has*
1) number of the distinct solutions of (3.1) *equals* $r(S)$ *and* $\mathrm{nrs}\{(3.1)\} = \sigma(S)$*;*
2) number of the solutions of (3.1) *satisfying the condition* $G(x,y) = 0$ *equals* $N - r(C)$ *(in case if there exists a solution of* (3.1) *of multiplicity* m_1 *that is, at the same time, a zero for* $G(x,y)$ *of multiplicity* m_2 *it should be counted* $\min(m_1, m_2)$ *times);*
3) if the matrix H *constructed by* (3.12) *is nonsingular, then all the solutions of* (3.1) *are distinct and the formula* (3.11) *remains valid.*

PROOF. We shall present here only the idea of the proof for the part 3 and, simultaneously, for the Theorem 3.2. In both cases, the starting point is the quadratic form in real variables $x_0, \dots, x_{N-1}$:

$$\sum_{j=1}^{N} G(\alpha_j,\beta_j)[x_0\Psi_0(\alpha_j,\beta_j)+x_1\Psi_1(\alpha_j,\beta_j)+\ldots+x_{N-1}\Psi_{N-1}(\alpha_j,\beta_j)]^2 \qquad (3.14)$$

where the system of monomials $\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1}$ is linearly independent on the solutions of the system (3.1). Under Assumption 3.2, one may take $\Psi_k(\alpha,\beta):=\alpha^k$, and then the matrix of the form (3.14) coincides with (3.12). Under Assumption 3.2′, one should take the following system

$$\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1} := \left\{ \begin{array}{l} 1, \alpha, \alpha^2, \dots, \alpha^{2n-4}, \alpha^{2n-3}, \alpha^{2n-2} \\ \beta, \alpha\beta, \alpha^2\beta, \dots, \alpha^{2n-4}\beta, \\ \beta^2, \alpha\beta^2, \dots, \alpha^{2n-6}\beta^2, \\ \dots, \\ \beta^{n-1} \end{array} \right\} \qquad (3.15)$$

Let us prove that

$$(\det[\Psi_{k-1}(\alpha_j,\beta_j)]_{k,j=1}^N)^2 = \det D \prod_{1\le j\le N} \mathcal{J}(\alpha_j,\beta_j) \qquad (3.16)$$

where $N\times N$–matrix D is made of the coefficients of the expansion (3.6) similarly to the matrix (3.13). Suppose first, that $\mathcal{J}(\alpha_j,\beta_j)\ne 0$ for all $j=\overline{1,N}$. The left hand side of (3.16) can be represented as

$$\prod_{1\le j\le N} \mathcal{J}(\alpha_j,\beta_j)\det\{[\Psi_{k-1}(\alpha_j,\beta_j)]_{k,j=1}^N([\Psi_{k-1}(\alpha_j,\beta_j)/\mathcal{J}(\alpha_j,\beta_j)]_{k,j=1}^N)^t\}$$

where ^{t} denotes the transposition of the matrix. By the equality (3.10), one has then the validity of (3.16). Thus, it is true under the assumption when all the solutions of (3.1) are simple. The set of the systems (3.1) possessing a multiple solution is of measure zero in the space of the coefficients of F_1 and F_2. The equality (3.16) is an algebraic one with respect to those coefficients. Therefore, it will be true also in general case.

Due to the equality (3.8), matrix D may be put into the block–triangular form by means of some elementary transformations. The blocks on the diagonal turn out to be of the Hankel type

$$\det[d_{2n-j-k-2,j+k}]_{j,k=0}^{p-1} = \mathcal{R}_p(A_0)/A_0 \text{ if } p=\overline{1,n-1},$$

$$\det[d_{2n-j-k-2,j+k}]_{j,k=0}^{n-1} = 1/A_0 \qquad (3.17)$$

So, under Assumption 3.2′, the determinant in the left hand side of (3.16) does not vanish, and the monomial system (3.15) is linearly independent on the solutions of (3.1).

Formula (3.11) then follows from the inertia law for quadratic forms.□

COROLLARY 3.2. *One has*

$$\det S = \kappa \prod_{1\le j\le N}\mathcal{J}(\alpha_j,\beta_j), \det C = \kappa \prod_{1\le j\le N} G(\alpha_j,\beta_j),$$

$$\det H = \det S \prod_{1\le j\le N} G(\alpha_j,\beta_j)$$

where $\kappa = (-1)^{(n-1)(n-2)/2}[\mathcal{R}_1(A_0)\ldots\mathcal{R}_{n-1}(A_0)]^2/A_0^{2n-1}$.

REMARK. As a matter of fact, the only essential assumption in the above theorem is Assumption 3.1, whereas Assumption 3.2′ is not. The structure of a monomial basis $\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1}$ is determined from the position of the nonzero Hankel minors of the determinant (3.17). We chose the basis (3.15) only because of its elegant relationship with subresultants.

Let us consider now the problem of finding the $\max f(x,y)$, where $f(x,y)$ is a real polynomial of a degree $n+1$.

On eliminating the trivial case $\max f(x,y) = +\infty$, we shall restrict ourselves to the case when the expansion of $f(x,y)$, in decreasing powers of x and y

$$f(x,y) := f_{(n+1)}(x,y) + f_{(n)}(x,y) + \cdots + f_{(0)}$$

begins with the homogeneous polynomial

$$f_{(n+1)} := a_{n+1,0}x^{n+1} + a_{n,1}x^n y + \ldots + a_{0,n+1}y^{n+1}$$

of an even degree $n+1$, and,moreover, $F(y) := f_{(n+1)}(1,y) < 0$ for every $y \in \mathbb{R}$. The latter condition holds iff $a_{0,n+1} < 0$ and $F(y)$ does not have real roots; thus, it can be verified by Theorem 2.1.

Consider the system

$$\frac{\partial f}{\partial x} = 0, \frac{\partial f}{\partial y} = 0 \tag{3.18}$$

yielding the critical points of $f(x,y)$. We shall suppose, that the features of $f(x,y)$ guarantee us the fulfilment of Assumption 3.1 and either of Assumptions 3.2 or 3.2′ for the system (3.18).

As in the previous section, let us consider

PROBLEM 3.1. For every real z find

$$\operatorname{nrs}\{(3.18) | f(x,y) > z\} \tag{3.19}$$

By using Theorems 3.2 or 3.3 for the case $F_1 := \partial f/\partial x$, $F_2 := \partial f/\partial y$, and $G := f(x,y) - z$, one can solve the Problem 3.1 in exactly the same manner as the Problem 2.1. Calculating the leading principal minors of the matrices (3.12) or (3.13) we get the the polynomial sequence in z :

$$H_1(z), \ldots, H_N(z) \ (N = n^2 \ ). \tag{3.20}$$

By Corollaries 3.1 and 3.2, the polynomial $H_N(z)$ has the roots which coincide with the critical values of $f(x,y)$. By multiplying it by the corresponding constant we get the polynomial

$$\mathcal{F}(z) := (-1)^N A_0^n \prod_{1 \le j \le N} (f(\alpha_j,\beta_j) - z) \tag{3.21}$$

the coefficients of which are polynomials with respect to the coefficients of $f(x,y)$, and its constant term equals $(-1)^N \mathcal{D}(f(x,y))$ (Schläfli (1953)).

ASSUMPTION 3.3. Let $\mathcal{D}(\mathcal{F}(z)) \neq 0$.

Under this assumption , there exists a one-to-one correspondence between the real solutions of (3.18) and those of $\mathcal{F}(z)$. Therefore the sequence (3.20) turns out to be the Sturm series for $\mathcal{F}(z)$, i.e.

$$\begin{aligned} &\operatorname{nrr}\{\mathcal{F}(z) = 0 | a < z < b\} = \operatorname{nrs}\{(3.18) | a < f(x,y) < b\} \\ &= P(1, H_1(a), \ldots, H_N(a)) - P(1, H_1(b), \ldots, H_N(b)), \end{aligned} \tag{3.22}$$

where P is the number of permanences in the corresponding sequence.

Once the greatest (real) root of $\mathcal{F}(z)$ is found, one can get the coordinates of the corresponding critical point (α, β) of $f(x, y)$ as the rational functions of it. Thus, for example, if the sequence (3.20) was constructed according with Theorem 3.2 (i.e. , for the matrix (3.11)) , then the x–component of the critical point may be evaluated by the formulae (2.9), where one should put $n := N, c_k := h_{k,0}$, while A_0, A_1 stand for the coefficients of x^N and x^{N-1} in $\mathcal{X}(x)$. As for the case when the series (3.20) is constructed according with Theorem 3.3, it will be illustrated by the following

EXAMPLE 3.1 Find the $\max f$ for

$$f(x,y) := -3x^4 + 24x^3y - 72x^2y^2 + 72xy^3 - 24y^4 - 4x^3 - 24x^2y + 72xy^2$$
$$-56y^3 + 30x^2 + 48xy + 180x + 96y$$

To solve the example, let us use Theorem 3.3. To find the elements of (3.13), it is convenient to evaluate first $s_{jk} := \sum_{p=1}^{9} \alpha_p^j \beta_p^k$ (analogues of the Newton sums for the system (3.18)):

$$s_{00} = 9, s_{1,0} = -81/5, s_{0,1} = -69/5, s_{2,0} = 7481/25, s_{1,1} = 4309/25, \dots$$

The sequence (3.20) for the matrix (3.13):

$$\begin{aligned}
H_1 &= -1125z + 662821,\\
H_2 &= 2136375z^2 - 4092726186z + 870894052211,\\
H_3 &= -70824375z^3 + 298756528245z^2 - 323525082079429z + 56416339849943559,\\
H_4 &= 3377997320625z^4 + \dots,\\
H_5 &= -144731302250625z^5 + \dots,\\
H_6 &= 207139392156375z^6 + \dots,\\
H_7 &= -1578710683329750z^7 + \dots,\\
H_8 &= 15(210901616449875z^8 - 1119967727621090904z^7 + 1810545656904280138068z^6\\
&\quad -973533633938806945166568z^5 + 151118470890640776270051685 0z^4\\
&\quad +157512964193185316621789948 40z^3 - 3086186500985790783485433998892z^2\\
&\quad -29207728791096637162171056853832 8z - 553907504350506277099645454062001 3),\\
H_9 &= -2460375z^9 + 13046305743z^8 - 20953332885564z^7 + 10858379628617100z^6\\
&\quad -1199221437495632850z^5 - 369773782407882562734z^4\\
&\quad +33574934405487787787124z^3 + 836331012136118485006470 0z^2\\
&\quad +43870230876294064609439652 9z + 66726857764901884700565619 43.
\end{aligned}$$

All the minors, except for H_8, were divided by positive rationals, and, up to those numbers, the leading coefficients of $H_k(z)$ are the leading principal minors for the matrix S from Theorem 3.3. According to it, all the critical points of $f(x, y)$ are real and distinct. Using the formulae (3.22), one can separate the roots of $H_9(z)$ and find out that its maximal root (and so, $\max f$) equals 2797.86763 ± 10^{-5}. To evaluate now the corresponding critical point (α, β) let us do the following procedure. Construct the matrix $\tilde{H}(z, s, t)$ having the same structure as (3.13) with the elements

$$\tilde{h}_{kl}(z) := st\ h_{kl}(z) - s\ h_{k,l+1}(z) - t\ h_{k+1,l}(z)$$

Expand its leading principal minor $\tilde{H}_8$ in decreasing powers of s and t. One has

$$\tilde{H}_8 \equiv s^8t^8H_8(z) + s^8t^7\mathrm{B}(z) + s^7t^8\mathrm{A}(z) + \dots$$

where

$$\begin{aligned}
\mathrm{B}(z) := {} & 2(20146311517771125z^8 - 106570998612435327624z^7 \\
& +170872213575475231298028z^6 - 90032548034190313784324088z^5 \\
& +1336867510460572526207165379\,0z^4 + 149853294415745316040066944660\,0z^3 \\
& -26569021824508842273514599780757\,2z^2 \\
& -25737131777190869121376463754566088z \\
& -4907004028522421107568929639028026\,03); \\
\mathrm{A}(z) := {} & 768(60696354196125z^8 - 320712403510525149z^7 \\
& +512978144492217304593z^6 - 268651356560467406295453z^5 \\
& +393364421525087195617342 35z^4 + 450952145771476685527021 7025z^3 \\
& -774957680752023422284994807757z^2 - 756510198738153495866666874738 63z \\
& -144513457949403109735240806622 8028).
\end{aligned}$$

If z is a root of $H_9(z)$, then the corresponding (α, β) can be found by the formulae

$$\alpha = s_{1,0} + \mathrm{A}(z)/H_8(z); \;\; \beta = s_{0,1} + \mathrm{B}(z)/H_8(z)$$

After substituting the value of $\max f$ in them, we get the critical point:

$$(-8.07285 \pm 10^{-5}, -11.50294 \pm 10^{-5}).$$

To conclude this section, we would like to mention the following: Theorems 2.3 and 3.3 make a basis for the recurrent method of the resultant calculation. In fact, coefficients c_j of the expansion (2.1) were the symmetric functions of the roots of a polynomial $F(x)$; the resultant $\mathcal{R}(F, G)$ can be then determined as the determinant of the *Hankel* matrix C (2.4), made of those coefficients (v. Corollary 2.3). Since we can now find such a resultant, let us calculate $\mathcal{X}(x)$ and $\mathcal{Y}(y)$ for the system (3.1) and make the expansion (3.9) for the fraction $G(x, y)\mathcal{V}(x, y)/[\mathcal{X}\mathcal{Y}]$. Coefficients of that expansion turn out to be the symmetric functions of solutions of the system (3.1). Now the resultant of three polynomials of two variables $\mathcal{R}(F_1(x, y), F_2(x, y), G(x, y))$ can be determined as the determinant of the block–*Hankel* matrix C from the Theorem 3.3 (v. Corollary 3.2). Since we can find now the resultant of three polynomials of two variables, we are able to extend in $\mathbb{R}^3$ the Jacobi method for finding the symmetric functions of solutions of a system of three equations in three variables and, simultaneously, to make an analogue in $\mathbb{R}^3$ of the Hermite method for separation of solutions of that system, etc. The leading principal minors of the matrices C play the role of subresultants.

To extend the method to $\mathbb{R}^3$, we will be in need of the linear representation of the resultant $\mathcal{R}(F_1, F_2, G)$. This will be given by the following

THEOREM 3.4 *Let, again, be* $deg F_1 = deg F_2 = n, m := deg G \le n$ *and* $C_N \neq 0$. *Suppose also that the* Assumption 3.2$'$ *is fulfiled. There exists a single set of polynomials* $\mathcal{M}_1(x, y), \mathcal{M}_2(x, y)$ *and* $\mathcal{Q}_1(x, y)$ *of the degrees satisfying the following restrictions*

$$\left\{ \begin{array}{l} deg\mathcal{M}_1 = deg_x\mathcal{M}_1 \le n + m - 2, \quad deg_y\mathcal{M}_1 \le n - 1; \\ \\ deg\mathcal{M}_2 = deg_x\mathcal{M}_2 \le n + m - 2, \quad deg_y\mathcal{M}_2 \le n - 1; \\ \\ deg\mathcal{Q}_1 = deg_x\mathcal{Q}_1 \le 2n - 2, \quad deg_y\mathcal{Q}_1 \le n - 1 \end{array} \right. \tag{3.23}$$

that provide the fulfilment of the identity

$$\mathcal{M}_1(x,y)F_1(x,y)+\mathcal{M}_2(x,y)F_2(x,y)+\mathcal{Q}_1(x,y)G(x,y)\equiv\frac{C_N}{d_{2n-2,0}} \tag{3.24}$$

PROOF. Let us rearrange the rows and columns of the matrix C, ordering the monomials of the basis (3.15) according to the magnitude of their degrees.

For simplicity, denote the new matrix also by C. From the equalities (3.3) and (3.4)

$$\mathcal{M}F_1+\mathcal{N}F_2\equiv\mathcal{X},\quad \mathcal{P}F_1+\mathcal{Q}F_2\equiv\mathcal{Y}$$

let us express F_1 and F_2 : $F_1\equiv(\mathcal{X}\mathcal{Q}-\mathcal{N}\mathcal{Y})/\mathcal{V}, F_2\equiv(\mathcal{M}\mathcal{Y}-\mathcal{P}\mathcal{X})/\mathcal{V}$. Substitute those expressions in (3.24) :

$$M_x(x,y)\mathcal{X}(x)+M_y(x,y)\mathcal{Y}(y)+\mathcal{Q}_1(x,y)G(x,y)\mathcal{V}(x,y)\equiv(C_N/d_{2n-2,0})\mathcal{V}(x,y) \tag{3.25}$$

Here

$$M_x\equiv\mathcal{M}_1\mathcal{Q}-\mathcal{M}_2\mathcal{P},\quad M_y\equiv\mathcal{M}_2\mathcal{M}-\mathcal{M}_1\mathcal{N} \tag{3.26}$$

According to the restrictions (3.23), M_x and M_y have to satisfy the conditions

$$\left\{\begin{array}{l} deg M_x = deg_x M_x \le N+m-2,\quad deg_y M_x\le 2n-2;\\ \\ deg M_y\le N+m-2,\quad deg_x M_y\le 2n+m-3,\quad deg_y M_y\le N-1\end{array}\right. \tag{3.27}$$

Divide the equality (3.25) by the product $[\mathcal{X}(x)\mathcal{Y}(y)]$ and expand both its sides in the series of negative powers of x and y (using the expansions (3.6), (3.9), and also the condition (3.8)):

$$\begin{aligned}\frac{M_x(x,y)}{\mathcal{Y}(y)}+\frac{M_y(x,y)}{\mathcal{X}(x)}+\mathcal{Q}_1(x,y)[L_1(x,y)+\sum_{j,k=0}^{\infty}\frac{c_{jk}}{x^{j+1}y^{k+1}}]\\ \equiv(C_N/d_{2n-2,0})\sum_{\substack{j,k=0\\ j+k\ge 2n-2}}^{\infty}\frac{d_{jk}}{x^{j+1}y^{k+1}}\end{aligned} \tag{3.28}$$

For definiteness, consider the case $n=3$. Let us find $\mathcal{Q}_1(x,y)$ in the form

$$\mathcal{Q}_1(x,y)=[1,x,y,x^2,xy,y^2,x^3,x^2y,x^4]\mathbf{Q}_1^t$$

where $\mathbf{Q}_1=[q_{00},q_{10},q_{01},q_{20},q_{11},q_{02},q_{30},q_{21},q_{40}]$ is the coefficient vector. Compare coefficients of the terms

$$\frac{1}{xy},\frac{1}{x^2y},\frac{1}{xy^2},\frac{1}{x^3y},\frac{1}{x^2y^2},\frac{1}{xy^3},\frac{1}{x^4y},\frac{1}{x^3y^2},\frac{1}{x^5y}$$

in the equality (3.28). In the right hand side all of them vanish except for that one of $x^{-5}y^{-1}$, the latter equals $d_{4,0}$. So, one gets the linear system for determing the coefficients of $\mathcal{Q}_1(x,y)$:

$$C\mathbf{Q}_1^t = [\underbrace{0, 0, ..., 0}_{8}, C_9]^t$$

Solving it by Cramer's rule, we obtain the following result:

The polynomial $\mathcal{Q}_1(x, y)$ from the equality (3.24) equals the determinant of the matrix obtained on replacing the last row in the matrix C by the row of monomials of the basis (3.15) (where one should replace $\alpha \to x,\ \beta \to y$) arranged according to the magnitude of their degrees.

It is seen that the above algorithm for finding $\mathcal{Q}_1(x, y)$ is similar to that of finding the polynomial $q(x)$ from the linear representation of the resultant for the case of one variable (v. the proof of the Theorem 3.3).

To find now $M_y(x, y)$ let us exploit the equality (3.25). Equate the coefficients of $y^{2N-2}, ..., y^N$ in its both sides. Because of the restrictions on the degrees (3.7) and (3.27), only those coefficients differ from zero which contain in the summands $M_y\mathcal{Y}$ and $\mathcal{Q}_1 G\mathcal{V}$. By recurrent formulae, it is possible to find the coefficients of M_y as the polynomials in x. To find then M_x, one may use the equality (3.25).Finally, $\mathcal{M}_1(x, y)$ and $\mathcal{M}_2(x, y)$ can be found from (3.26).

Another method for finding $\mathcal{M}_1$ and $\mathcal{M}_2$ can be proposed after getting the expression for $\mathcal{Q}_1(x, y)$. Consider the equality (3.24) rewritten as

$$\mathcal{M}_1 F_1 + \mathcal{M}_2 F_2 \equiv (C_N/d_{2n-2,0} - \mathcal{Q}_1 G)$$

together with (3.4) : $\mathcal{M}F_1 + \mathcal{N}F_2 \equiv \mathcal{X}(x)$. Then one has

$$\mathcal{M}_1(x, y) := \frac{1}{\mathcal{X}(x)}[remainder\ on\ division\ of\ \mathcal{M}(C_N/d_{2n-2,0} - \mathcal{Q}_1 G)\ by\ F_2(x, y)],$$

$$\mathcal{M}_2(x, y) := \frac{1}{\mathcal{X}(x)}[remainder\ on\ division\ of\ \mathcal{N}(C_N/d_{2n-2,0} - \mathcal{Q}_1 G)\ by\ F_1(x, y)]$$

The functions in brackets are considered as polynomials in y. One may notice the analogy with the remark at the end of the proof of Theorem 2.3.$\square$

The background for the above mentioned 'shuttle' procedure (i.e. , symmetric functions of solutions of a polynomial system $F_1(x_1, \ldots, x_{n-1}) = 0, \ldots, F_{n-1}(x_1, \ldots, x_{n-1}) = 0 \Rightarrow$ resultant of the polynomials $F_1, \ldots, F_{n-1},\ F_n(x_1, \ldots, x_{n-1}) \Rightarrow$ symmetric functions of solutions of a polynomial system $\tilde{F}_1(x_1, \ldots, x_n) = 0, \ldots, \tilde{F}_n(x_1, \ldots, x_n) = 0 \Rightarrow \ldots$) can be found in the 19th century (Schläfli (1953)).The idea of constructing the resultant as the block-Hankel matrix can be found in (Laurent (1900),p.42); though there were several errors mentioned in (Macaulay (1903),p.4). As for the relationships of the proposed *Hankel* approach for the resultant computation with the others existed (for example, the determinantial one - Gonzalez-Vega (1991)), as well as with the Shape lemma from the Gröbner basis theory, we hope to discuss them in subsequent papers.

4. The constrained maximum

More interesting is, of course, the problem of searching for constrained maximum, i.e.

$$\max_{X\in\mathbb{S}} f(X), \ \ X := (x_1,\dots,x_k), \ \ \deg f := n+1$$

where $\mathbb{S}$ is the set of real solutions of the system of algebraic inequalities (constraint set)

$$G_1(X)\ge 0,\dots,G_p(X)\ge 0 \tag{4.1}$$

In the computer algebra that problem can be solved, for example, by using CAD algorithm applied to the set $\{X\in\mathbb{R}^k | f(X)-z\ge 0, G_1(X)\ge 0,\dots,G_p(X)\ge 0\}$ (Bank, Heintz, Krick, Mandel and Solernó (1992)).

In the present section we shall apply the Hermite method described in the previous sections to the stated problem. We restrict ourselves to the case of two variables ($k=2$). We will need the following formulae

$$\begin{aligned}&\operatorname{nrs}\{(3.1)|G_1>0,\dots,G_p>0\}\\ &=\frac{1}{2^{p-1}}[(1-2^{p-1})\operatorname{nrs}\{(3.1)\}+\sum_{1\le j_1\le p}\operatorname{nrs}\{(3.1)|G_{j_1}>0\}\\ &+\sum_{1\le j_1<j_2\le p}\operatorname{nrs}\{(3.1)|G_{j_1}G_{j_2}>0\}+\dots+\operatorname{nrs}\{(3.1)|G_1G_2\dots G_p>0\}]\end{aligned} \tag{4.2}$$

established apparently by Markov (Markoff(1940)). Though the well–known results by Ben–Or, Kozen and Reif (1986) permit one to compute that number in a more efficient way, formulae (4.2) will be enough to explain the idea of our approach.

Consider, first, the case when $p=1$, i.e. we look for $\max_{G(x,y)\ge 0} f(x,y)$.

Assumption 4.1. We shall assume the domain $G(x,y)\ge 0$ to be bounded. This can be achieved by imposing on polynomial G the restrictions similar to those made in Section 3 for the function f (in particular, let $m := \deg G$ be even).

Function f attains its maximum either at a critical point lying in the interior of the constraint set, or on its boundary $G=0$. The latter is possible iff the point (of relative maximum) satisfies the polynomial system

$$G(x,y)=0, \mathcal{J}(G,f) := \partial G/\partial x \partial f/\partial y - \partial f/\partial x \partial G/\partial y = 0 \tag{4.3}$$

Similar to Section 3 consider then

Problem 4.1. For every real z find

$$\operatorname{nrs}\{(3.18)|G>0, f>z\} \quad \text{and} \quad \operatorname{nrs}\{(4.3)|f>z\}$$

To compute the number $\operatorname{nrs}\{(3.18)|f>z\}$ we have constructed in Section 3 the polynomial sequence (3.20):

$$H_1(z),\dots,H_{n^2}(z)$$

Critical values of f coincide with the roots of polynomial $H_{n^2}(z)$, or with those of $\mathcal{F}(z)$ given by (3.21). The overall sequence is the Sturm series for $\mathcal{F}(z)$. In exactly the same manner, one can find the similar polynomial sequence

$$\hat{H}_1(z),\dots,\hat{H}_{(m-1)mn}(z) \tag{4.4}$$

for the number $\mathrm{nrs}\{(4.3)|f > z\}$ (under assumptions of Theorem 3.2 or 3.3). By Corollaries 3.1 and 3.2, the roots of $\hat{H}_{(m-1)mn}(z)$ coincide with the values of f at the solutions of (4.3). The sequence (4.4) turns out to be the Sturm series for the problem of separation of those roots.

To find now the $\mathrm{nrs}\{(3.18)|G > 0, f > z\}$ one can either apply directly the formulae (4.2) and construct the third Sturm series for the number $\mathrm{nrs}\{(3.18)|(f-z)G > 0\}$; or to transform the inequality $G(x,y) > 0$ into a new one with respect to the variable z. The latter is possible since the roots of $\mathcal{F}(z)$ and the coordinates of corresponding critical points are connected by the rational functions:

$$x = r_1(z), y = r_2(z)$$

(v. Example 3.1). Denote $\mathcal{G}(z) := G(r_1(z), r_2(z))$, then

$$\mathrm{nrs}\{(3.18)|G(x,y) > 0, b > f(x,y) > a\} = \mathrm{nrr}\{\mathcal{F}(z) = 0|\mathcal{G}(z) > 0, b > z > a\}.$$

In this way, Problem 4.1 is reduced now to a similar one for the single variable. Though it could be solved purely algebraically, it looks more interesting to apply hybrid methods. For example,by using the systems (3.20) and (4.4) one can separate the real roots of $\mathcal{F}(z)$ and $\hat{H}_{(m-1)mn}(z)$ by rational numbers and arrange them (checking, if necessary, that $\mathcal{R}(\mathcal{F}(z), \hat{H}_{(m-1)mn}(z))$ does not vanish). If the maximal of those roots is that one of $\hat{H}_{(m-1)mn}(z)$, then $\max_{G \geq 0} f$ is attained at the boundary of the constraint set. As for the case, when the maximal root belongs to $\mathcal{F}(z)$, one has to check the rational inequality $\mathcal{G} > 0$ on it. That can be done using the interval arithmetic (Collins and Krandick (1993)).

The proposed approach can be extended to the general case when the constraint set $\mathbb{S}$ is given by the system of inequalities (4.1). In addition to the systems (3.18) and (4.3) one other type of systems appears: $G_p(x,y) = 0, G_q(x,y) = 0$ yielding the 'corner points' of $\mathbb{S}$.

To conclude this section we would like to mention the following: the problem of consistency of the system of inequalities (4.1) can be considered itself as the problem of constrained maximum (Uteshev and Cherkasov (1995)). Thus, for example, under Assumption 4.1, the set defined by $G(x,y) \geq 0$ is not empty iff $\mathrm{nrs}\{\partial G/\partial x = \partial G/\partial y = 0|G > 0\}$ is greater then zero. Moreover, on this way one can even establish the geometry of its boundary: the difference between the number of positive and the number of negative ovals of $G = 0$ equals

$$\mathrm{nrs}\{\partial G/\partial x = \partial G/\partial y = 0|G > 0, \mathcal{H}(G) > 0\}$$

$$-\mathrm{nrs}\{\partial G/\partial x = \partial G/\partial y = 0|G > 0, \mathcal{H}(G) < 0\}$$

Here $\mathcal{H}(G) := (\partial^2 G/\partial x^2)(\partial^2 G/\partial y^2) - (\partial^2 G/\partial x \partial y)^2$ is the Hessian of G (Petrowsky (1938)).

5. Conclusions.

We have treated here the problem of finding the maximum value of a polynomial in an algebraic domain. The proposed approach is based on Hermite's ideas for the problem of separation of solutions of algebraic systems. We have also discussed the possibility of making the separation/elimination algorithm self–contained (only using Hankel or block–Hankel matrices).

Several problems remain unsolved, including the following :
1. Establishing the structure of Assumptions 2.2 and 3.3 in terms of the coefficients of f (maybe, via invariants).
2. Comparing the elimination procedure outlined at the end of Section 3 with others existing, and estimating its computational complexity.

References

Bank B., Heintz J., Krick T., Mandel R., Solernó P.(1992) Computability and complexity of polynomial optimization problems, in *Modern Methods of Optimization* (Krabs W., Zowe J. eds.) Springer, 1–23

Becker E., Wörmann T. (1994). On the trace formula for quadratic forms. *Contemp. Math.* 155:271-291

Ben–Or M.,Kozen D.,Reif J.(1986).The complexity of elementary algebra and geometry. *J. of Comp. and Sys. Sciences* 32:251-264

Collins G.E. (1971). The calculation of multivariate polynomial resultants. *J. of the ACM* 18(4):515-532.

Collins G.E., Krandick W. (1993). A hybrid method for high precision calculation of polynomial real roots. *Proceedings of International Symposium on Symbolic and Algebraic Computation, Kiev 1993* (M.Bronstein, ed.) ACM Press, 47-52

Gantmacher F.R.(1959) *The Theory of Matrices.* New York: Chelsea

Gonzalez-Vega L. (1991) Determinantal formulae for the solution set of zero–dimensional ideals. *J. Pure Appl. Algebra* 76:57-80

Gonzalez-Vega L., Lombardi H., Recio T., Roy M.-F. (1990) Specialization de la suite de Sturm et sous-resultants. *Inform. Theor. et Applic.* 24(6):561-588

Hermite Ch. (1912) Sur l'extesion du théoreme de M.Sturm a un système d'équations simultannées. *Oeuvres* 3:1–34, Paris: Gauthier–Villars

Krein M.G., Naimark M.A.(1981) The method of symmetric and Hermitian forms in the theory of separation of the roots of algebraic equations. *Linear Multilin. Algebra* 10(4):265–308

Kronecker L.(1897) Zur Theorie der Elimination einer Variabeln aus zwei algebraischen Gleichungen . *Werke* 2:113–192 , Leipzig: Teubner

Laurent H. (1900) *L'Élimination* , in *Scientia, Phys.-Mathématique*,7, Paris:Gauthier–Villars

Macaulay F.S.(1903). On some formulae in elimination. *Proc. London Math. Soc., Ser.1*, 35:3–27

Markoff A. (1940). On the determination of the number of roots of an algebraic equation situated in a given domain. *Mat.Sb.* 7(1):3-6

Pedersen P., Roy M.-F., Szpirglas A.(1993) Counting real zeros in the multivariate case. *Progress Math.* 109, 203–223

Petrowsky I. (1938). On the topology of real plane algebraic curves. *Ann.Math.* 39(1):189-209

Schläfli L. (1953) Über die Resultante eines Systemes mehrerer algebraischer Gleichungen. *Gesammelte Math. Abhandlungen* 2:9–112, Basel: Birkhäuser

Uteshev A.Yu., Shulyak S.G.(1992) Hermite's method of separation of solutions of systems of algebraic equations and its applications, *Linear Algebra Appl.* 177:49–88

Uteshev A.Yu.,Cherkasov T.M.(1995) Localization of solutions of systems of algebraic equations and inequalities – Hermite's method. Accepted for publication in *Doklady AN Rosii* (*Russian Math. Doklady*.