J. Symbolic Computation (1998) **25**, 587–618

The Search for the Maximum of a Polynomial

ALEXEI YU. UTESHEV† AND TIMOFEI M. CHERKASOV‡

Faculty of Applied Mathematics, St Petersburg State University, Bibliotechnaja pl.2, Petrodvoretc, 198904, St Petersburg, Russia

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 the univariate polynomial $\mathcal{F}(z) := \prod_j (z - f(\Lambda_j))$, where the product is extended over all the critical points Λ_j of $f(X)$. Employment of the Hermite method of separation of real solutions of an algebraic equation system permits one to construct along with $\mathcal{F}(z)$ its Sturm series, and to restore the coordinates of the 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 is devoted to the nonlinear optimization problem: finding the maximum of a polynomial function in the domain defined by a polynomial inequality system. This polynomial programming problem has been intensively treated by several authors during the last two decades; for a survey of the approaches, which exploit essentially the polynomiality of the stated problem, we refer to Bank *et al.* (1992).

We are concerned here mainly with the problem of finding

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

for polynomial functions f and G.

The idea of the proposed approach can be described briefly as a reversion of the usual algorithm

$$\text{critical points} \;\rightarrow\; \text{critical values}. \tag{1.1}$$

We intend to construct first an univariate polynomial $\mathcal{F}(z)$ whose roots coincide with the critical values of the objective function $f(X)$. Due to the polynomiality of the problem, this can be done purely algebraically, i.e. by means of the finite number of elementary algebraic operations on the coefficients of f. The problem can be reduced to the one of finding the multivariate resultant, i.e. a polynomial function

$$\mathcal{R}(F_1(X), \ldots, F_K(X), G(X)) \tag{1.2}$$

† E-mail: irina@utesh.spb.su
‡ E-mail: tcherk@medlib.spb.ru

0747–7171/98/050587 + 32 $25.00/0 sy970190

in the coefficients of the given polynomials $F_1(X), \dots, F_K(X), G(X) \in \mathbb{C}[X]$ whose vanishing (under certain assumptions) gives a necessary and sufficient condition for the existence of a common zero for these polynomials. We define $\mathcal{F}(z)$ to be

$$\mathcal{F}(z) := \mathcal{R}(\partial f/\partial x_1, \dots, \partial f/\partial x_K, f(X) - z).$$

The resultant can be computed with the help of either some determinantal representations (like Cayley's or Macaulay's ones (Macaulay, 1903, 1916)) or via the Gröbner basis construction.

Let us take into account, however, the fact that finding the polynomial $\mathcal{F}(z)$ is not the final aim of the optimization problem. The next task is to separate the real roots of $\mathcal{F}(z)$. For a univariate polynomial, this problem is usually solved via the Sturm series construction. Our suggestion is to do this simultaneously with $\mathcal{F}(z)$. For that aim, one can use the method for the separation of the solutions of a system of algebraic equations discovered by Charles Hermite in 1852–1856 (Hermite, 1912), and developed by Francesco Brioschi and other scholars in the 19th century (a historical review can be found in Krein and Naimark (1981) and Uteshev and Shulyak (1992)). Recently the method has been revised by several authors: Becker and Wörmann (1994) and Pedersen *et al.* (1993); Uteshev and Shulyak (1992) and Uteshev and Cherkasov (1996). The general separation problem is formulated as follows: Find the number of real solutions of the polynomial system

$$F_1(X) = 0, \dots, F_K(X) = 0 \tag{1.3}$$

which satisfy the polynomial inequality $G(X) > 0$, i.e.

$$\mathrm{nrs}\{(1.3) \mid G(X) > 0\}. \tag{1.4}$$

Here $\{F_1(X), \dots, F_K(X), G(X)\} \subset \mathbb{R}[X]$. Briefly, the idea of Hermite's method consists of finding the rank and signature of the quadratic form in real $x_0, \dots, x_{N-1}$:

$$\sum_{j=1}^{N} G(\Lambda_j)[x_0\Psi_0(\Lambda_j) + \cdots + x_{N-1}\Psi_{N-1}(\Lambda_j)]^2. \tag{1.5}$$

Here $\Lambda_1, \dots, \Lambda_N$ are the solutions of (1.3), and the system of monomials $\{\Psi_j(\Lambda)\}_{j=0}^{N-1}$ is linearly independent on these solutions. The elements of the matrix H of the form (1.5) are the symmetric polynomials of the solutions. They can be expressed rationally in terms of the coefficients of the polynomials $F_1, \dots, F_k, G$. Hence, Hermite's method permits one to find the number (1.4) purely algebraically, and this fact can be utilized for the optimization problem. We state the latter as the one of finding

$$\mathrm{nrs}\{\partial f/\partial x_1 = 0, \dots, \partial f/\partial x_K = 0 \mid f(X) > z\} \tag{1.6}$$

or

$$\mathrm{nrs}\{\partial f/\partial x_1 = 0, \dots, \partial f/\partial x_K = 0 \mid f(X) > z, G(X) > 0\}.$$

This statement gives us a perfect opportunity to reconsider the foundations of Hermite's method. Although it was sufficiently illuminated in the cited papers (and several others), there exist some reasons to do this. In most of them the desire is to justify the end by the means. Most of the papers on the method use, as their essential parts, the elements of the "outer" theories: e.g. Macaulay's determinantal representation for the resultant or the Gröbner basis construction for evaluating symmetric polynomials of solutions. We intend to avoid this and to make the method self-contained, i.e. to unify

the separation and the elimination process. We will introduce a new matrix C with a structure similar to that of the matrix H with the deficiency index (= order − rank) equal to the number of solutions of the system (1.3) common with $G(X)$. Its determinant differs from the resultant (1.2) only by the known factor.

We will treat in detail the cases of $K = 1$ and $K = 2$ variables, and show how to manage the extension to the case of $K = 3$ variables. This methodology permits us to clarify the ideas and to illustrate the difficulties with examples. If a cited classical result is contained in a rarely available publication, we will provide a sketch of the proof.

The paper is organized as follows. Section 2 contains some preliminary results from classical Elimination Theory: resultant, subresultants, discriminant for the univariate polynomials; eliminants, symmetric polynomials of solutions and Poisson's inductive definition of the resultant for the multivariate polynomials. Some assumptions are made with most of them guaranteeing the existence of exactly $N := \deg F_1 \dots \deg F_K$ (the Bézout number) solutions for the system (1.3).

In Section 3 we consider the univariate case. Employing results by Jacobi, Hermite, Sylvester and Kronecker, we construct the Hankel matrix with the determinant coinciding (up to a constant factor) with the $\mathcal{F}(z)$, and with the sequence of the leading principal minors playing the role of Sturm series for that polynomial.

Section 4 is devoted to the bivariate case. Jacobi's method is employed to evaluate the symmetric functions of solutions of the system (1.3). The two methods for choosing the monomial system $\{\Psi_j(\Lambda)\}_{j=0}^{N-1}$ in (1.5) are discussed. The general results are exploited to find the number (1.6) and to separate the roots of $\mathcal{F}(z)$ in exactly the same manner as for the univariate case.

In both sections, we mention an important consequence of the results on the existence of the zero common to the polynomials considered. Provided this zero is unique, it can be represented as a rational function of the coefficients of these polynomials. For the optimization problem, this implies the completion of the reversion of the scheme (1.1): once the maximum of $f(X)$ is found, one can restore the corresponding critical point. We also discuss the problem of the existence of "extraneous" real roots for $\mathcal{F}(z)$, i.e. those corresponding to the imaginary critical points.

As its essential part, Jacobi's method utilizes the polynomials $\mathcal{P}_1(X), \dots, \mathcal{P}_K(X)$, $\mathcal{Q}(X) \in \mathbb{C}[X]$ providing a linear representation for the resultant

$$\mathcal{P}_1 \cdot F_1 + \dots + \mathcal{P}_K \cdot F_K + \mathcal{Q} \cdot G \equiv \mathcal{R}(F_1, \dots, F_K, G).$$

For $K = 2$, we give an algorithm to find them in Section 5.

Section 6 is devoted to the constrained maximization problem. We develop a general approach: to transform the problem into one of the separation of the real roots of a univariate polynomial. Related topics are also discussed, such as establishing the topological structure of the constraint set.

2. Poisson's Definition of the Resultant

Let $X = (x_1, \dots, x_K)$ be a vector of variables. Consider polynomials $F_1(X), \dots, F_K(X), G(X) \in \mathbb{C}[X]$. Let $\deg F_j = n_j$, $(j = 1, \dots, K), \deg G = m$. In order to establish a necessary and sufficient condition for the existence of a common zero for $F_1, \dots, F_K$ and G over $\mathbb{C}$, we first recall the inductive definition of the resultant according to Poisson (Netto, 1896-1900; Schläfli, 1953). Although the resultant is usually defined for forms (homogeneous polynomials) complete in all their terms and parametric

coefficients, we shall restrict ourselves to the non-homogeneous and *generic* case. This means that we will impose some restrictions on the coefficients of the leading forms of the considered polynomials under which it is possible to deduce a single condition for the existence of a common zero for these polynomials.

As a basis for induction, we take $K = 1$ and consider polynomials

$$F(x) := A_0 x^n + \cdots + A_n, \quad G(x) := B_0 x^m + \cdots + B_m \quad (A_0 \neq 0. B_0 \neq 0). \tag{2.1}$$

If we denote the roots of $F(x)$ by $\lambda_1, \ldots, \lambda_n$, then, according to the well-known result by Gauss, the value of any symmetric polynomial of these roots can be represented as a rational function of the coefficients of $F(x)$. For example, here are expressions for the *Newton sums*:

$$s_k := \sum_{j=1}^{n} \lambda_j^k = \begin{cases} n & \text{if } k = 0\ ; \\ -A_1/A_0 & \text{if } k = 1\ ; \\ -(A_1 s_{k-1} + A_2 s_{k-2} + \cdots + A_{k-1} s_1 + A_k k)/A_0 & \text{if } k < n\ ; \\ -(A_1 s_{k-1} + A_2 s_{k-2} + \cdots + A_n s_{k-n})/A_0 & \text{if } k \geq n. \end{cases} \tag{2.2}$$

The *resultant* $\mathcal{R}(F, G)$ is defined formally as the following symmetric polynomial:

$$\mathcal{R}(F, G) = A_0^m G(\lambda_1) \ldots G(\lambda_n) \tag{2.3}$$

while it can practically be found by any of the well-known methods, e.g., the Sylvester determinant (Jury, 1974; Akritas, 1989):

$$\mathcal{R}(F, G) = (-1)^{n(n-1)/2} \det \mathcal{U}, \tag{2.4}$$

where

$$\mathcal{U} = \left(\begin{array}{cccccccc} A_0 & A_1 & A_2 & \ldots & A_n & 0 & \ldots & 0 \\ 0 & A_0 & A_1 & A_2 & \ldots & A_n & \ldots & 0 \\ & & & \ldots & & & \ldots & \\ 0 & \ldots & 0 & A_0 & A_1 & A_2 & \ldots & A_n \\ 0 & \ldots & 0 & 0 & B_0 & B_1 & \ldots & B_m \\ 0 & \ldots & 0 & B_0 & B_1 & \ldots & B_m & 0 \\ & \ldots & & & \ldots & & & \\ B_0 & B_1 & \ldots & B_m & 0 & \ldots & & 0 \end{array} \right) \begin{array}{l} \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} m \text{ rows} \\ \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} n \text{ rows.} \end{array} \tag{2.5}$$

Furthermore, consider the minor of $\mathcal{U}$ of the order $m + n - 2k$ obtained on deleting the k first and the k last rows, and the k first and the k last columns of $\mathcal{U}$:

$$\mathcal{R}_k(F, G) = \left| \begin{array}{ccccccccc} A_0 & A_1 & A_2 & \ldots & \ldots & \ldots & \ldots & A_{n+m-2k-1} \\ 0 & A_0 & A_1 & A_2 & \ldots & \ldots & \ldots & A_{n+m-2k-2} \\ & & & \ldots & \ldots & & & \\ 0 & \ldots & 0 & A_0 & A_1 & A_2 & \ldots & A_{n-k} \\ 0 & \ldots & 0 & 0 & B_0 & B_1 & \ldots & B_{m-k} \\ 0 & \ldots & 0 & B_0 & B_1 & \ldots & B_{m-k} & B_{m-k+1} \\ & & & \ldots & \ldots & & & \\ B_0 & B_1 & \ldots & \ldots & \ldots & \ldots & \ldots & B_{n+m-2k-1} \end{array} \right| \begin{array}{l} \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} m-k \text{ rows} \\ \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} n-k \text{ rows} \end{array} \tag{2.6}$$

(we set $A_j := 0$ for $j > n$ and $B_k := 0$ for $k > m$). It is called the kth *subresultant* of $\mathcal{R}(F, G)$. For simplicity, we term $\mathcal{R}_0(F, G)$ for the $\det \mathcal{U}$, so: $\mathcal{R}(F, G) = (-1)^{n(n-1)/2} \mathcal{R}_0(F, G)$. One has the following fundamental property of subresultants:

$$\deg(\gcd(F, G)) = d \iff \mathcal{R}_0 = 0, \mathcal{R}_1 = 0, \ldots, \mathcal{R}_{d-1} = 0, \mathcal{R}_d \neq 0. \tag{2.7}$$

For the particular choice $G(x) := F'(x)$, the expression

$$\begin{aligned}\mathcal{D}(F) &:= \frac{1}{A_0}(-1)^{n(n-1)/2}\mathcal{R}(F,F') \\ &= \frac{1}{n^{n-2}}\mathcal{R}_0\Big(\sum_{j=1}^{n} jA_jx^{n-j}, \sum_{j=0}^{n-1}(n-j)A_jx^{n-j-1}\Big) \qquad (2.8)\\ &= A_0^{2n-2}\prod_{1\le k<j\le n}(\lambda_j-\lambda_k)^2 \qquad (2.9)\end{aligned}$$

is known as the *discriminant* of $F(x)$. We will refer to the kth subresultant of the resultant (2.8) as the kth *subdiscriminant* and denote it by $\mathcal{D}_k(F)$. The relation of $\mathcal{D}_k$ to the existence of multiple roots for $F(x)$ is evidently established from (2.7).

Consider now the case of $K = 2$ variables. Let us use the above construction of the univariate resultant to construct the resultant for three polynomials $F_1(x,y), F_2(x,y)$, $G(x,y)$ in two variables. Consider first the system:

$$F_1(x,y) = 0, \quad F_2(x,y) = 0 \qquad (2.10)$$

and take the leading forms from the expansions of F_1 and F_2 in decreasing powers of x and y:

$$F_{j,n_j}(x,y) := A_{j,0}x^{n_j} + A_{j,1}x^{n_j-1}y + \cdots + A_{j,n_j}y^{n_j}. \qquad (2.11)$$

Assumption 2.1. *Let $A_{j,0} \neq 0, A_{j,n_j} \neq 0$ for $j = 1, 2$.*

Construct the resultants of F_1 and F_2 on elimination of the variables y and x respectively:

$$\mathcal{X}(x) := \mathcal{R}_y(F_1,F_2), \quad \mathcal{Y}(y) := \mathcal{R}_x(F_1,F_2) \qquad (2.12)$$

(known also as the *eliminants*). One has

$$\mathcal{X}(x) = \mathcal{A}_0x^N + \text{ lower order terms, } \mathcal{Y}(y) = (-1)^N\mathcal{A}_0y^N + \text{ lower order terms.}$$

Here

$$N := n_1n_2, \quad \mathcal{A}_0 := \mathcal{R}(F_{1,n_1}(1,y), F_{2,n_2}(1,y)). \qquad (2.13)$$

Assumption 2.2. *Let $\mathcal{A}_0$ defined by* (2.13) *be different from zero.*

Under this assumption, the number of solutions of (2.10) (counted in accordance with their multiplicities) equals N (Bézout's well-known theorem). If $\Lambda_j := (\alpha_j, \beta_j)$ is a solution of (2.10), 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 (2.10).

Definition. Function $\Phi(x_1,y_1;x_2,y_2;\ldots;x_\ell,y_\ell) : \mathbb{C}^{2\ell} \to \mathbb{C}$ is called a *symmetric* function of the ℓ pairs of variables $(x_1,y_1),(x_2,y_2),\ldots,(x_\ell,y_\ell)$ if its value is unchanged when any of the pairs are interchanged:

$$\Phi(x_1,y_1;x_2,y_2;\ldots;x_\ell,y_\ell) \equiv \Phi(x_{j_1},y_{j_1};x_{j_2},y_{j_2};\ldots;x_{j_\ell},y_{j_\ell})$$

for the distinct $j_1,\ldots,j_\ell$.

THEOREM 2.1. (SCHLÄFLI, 1953) *Under Assumptions* 2.1 *and* 2.2*, the value of any symmetric polynomial* Φ *of the* N *pairs of variables on the solution* $\Lambda_1, \dots, \Lambda_N$ *of the system* (2.10) *is a rational function in the coefficients of* F_1, F_2*.*

We will give an idea of the proof. Schläfli first proved that any symmetric polynomial can be represented as a polynomial from the *elementary* symmetric polynomials, i.e. the expressions $\sum_{p=1}^{\ell} x_p^j y_p^k$. In terms of the system (2.10) this means that the problem stated can be reduced to the one of evaluating the generalized Newton sums:

$$s_{jk} := \sum_{p=1}^{N} \alpha_p^j \beta_p^k.$$

This can be performed via Poisson's method (Uteshev and Shulyak, 1992). Let us introduce two new variables u and t by the formula $t = x + uy$ (known also as the Liouville substitution, or the *u-substitution*). Replacing x in (2.10) by $t - uy$, we obtain the polynomial system

$$F_1(t - uy, y) = 0, \quad F_2(t - uy, y) = 0 \tag{2.14}$$

Considering here t and y as variables and u as a parameter, and constructing the resultant

$$\mathcal{T}(t, u) := \mathcal{R}_y(F_1(t - uy, y), F_2(t - uy, y) \equiv \mathcal{A}_0 t^N + \tau_1(u) t^{N-1} + \dots + \tau_N(u) \tag{2.15}$$

we eliminate y from (2.14). One has: $\mathcal{T}(t, 0) \equiv \mathcal{X}(t)$, $\deg \tau_j(u) \le j$ $(j = 1, \dots, N)$. The component t_j of a solution (t_j, y_j) of the system (2.14) is a root of $\mathcal{T}(t, u)$. Let us find the Newton sums for $\mathcal{T}(t, u)$ using the formulae (2.2):

$$s_k(u) = \sum_{p=1}^{N} t_p^k = T(\tau_1(u), \dots, \tau_N(u))$$

where T is a polynomial function in $\tau_1(u), \dots, \tau_N(u)$ and, thus, polynomial in u. Expand T in powers of u and remember that $t_j = \alpha_j + u\beta_j$, with (α_j, β_j) being a solution of (2.10):

$$s_k(u) = \sum_{p=1}^{N} (\alpha_p + u\beta_p)^k = T_{k0} + T_{k1} u + T_{k2} u^2 + \dots. \tag{2.16}$$

Since this is an identity in u, comparing coefficients gives

$$\binom{k}{j} \sum_{p=1}^{N} \alpha_p^{k-j} \beta_p^j = T_{kj}.$$

The latter formula permits one to calculate every sum s_{jk}.□

By Theorem 2.1, the symmetric polynomial of solutions $G(\alpha_1, \beta_1) \dots G(\alpha_N, \beta_N)$ can be expressed rationally in terms of the coefficients of F_1 and F_2. It can be proved that the expression

$$\mathcal{R}(F_1, F_2, G) := \mathcal{A}_0^m G(\alpha_1, \beta_1) \dots G(\alpha_N, \beta_N) \tag{2.17}$$

is a polynomial (with integer coefficients) with respect to the coefficients of F_1, F_2 and G. Under Assumptions 2.1 and 2.2, its vanishing gives a necessary and sufficient condition for the existence of a common zero for F_1, F_2 and G. Expression (2.17) is called the *resultant*

of the polynomials considered. Some properties of the resultant can be established, in particular that $\mathcal{R}(F_1, F_2, G)$ is independent (up to its sign) to the order of its arguments.

We are now able to proceed to the case of $K = 3$ variables. For the polynomials $F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)$ and $G(x, y, z)$ consider first the system:

$$F_1(x, y, z) = 0, \quad F_2(x, y, z) = 0, \quad F_3(x, y, z) = 0 \tag{2.18}$$

and take the leading form $F_{j,n_j}(x, y, z)$ from the expansion of F_j in decreasing powers of x, y and z. Assumption 2.1 is replaced by:

ASSUMPTION 2.1′. *Let the polynomials* $F_{j,n_j}(0, y, z), (j = 1, 2, 3)$ *have no common zero.*

According to the previous step of the inductive process, this property can be verified via the univariate resultant (2.5). With the help of the bivariate resultant (2.17), one is also able to compute the eliminants, i.e. the resultants of the polynomials F_1, F_2, F_3 on eliminating pairs of variables. For example:

$$\mathcal{X}(x) := \mathcal{R}_{y,z}(F_1, F_2, F_3) = \mathcal{A}_0 x^N + \text{ lower order terms},$$

where

$$N := n_1 n_2 n_3, \quad \mathcal{A}_0 := \mathcal{R}(F_{1,n_1}(1, y, z), F_{2,n_2}(1, y, z), F_{3,n_3}(1, y, z)). \tag{2.19}$$

ASSUMPTION 2.2′. *Let* $\mathcal{A}_0$ *defined by* (2.19) *be different from zero.*

Under Assumptions 2.1′ and 2.2′, there exist exactly $N := n_1 n_2 n_3$ solutions $(\alpha_j, \beta_j, \gamma_j)$ for the system (2.18), with their x-components being the roots of $\mathcal{X}(x)$. One can then prove an analogue of Theorem 2.1 (using the u-substitution in the form $t = x+u_1 y+u_2 z$), and represent any symmetric polynomial of the solutions as a rational function of the coefficients of F_1, F_2 and F_3. In particular, the expression

$$\mathcal{R}(F_1, F_2, F_3, G) := \mathcal{A}_0^m G(\alpha_1, \beta_1, \gamma_1) \dots G(\alpha_N, \beta_N, \gamma_N) \tag{2.20}$$

can be represented as a polynomial (with integer coefficients) with respect to the coefficients of F_1, F_2, F_3 and G. It is called the resultant of the polynomials considered.

The procedure goes further in a similar way. So it can be described as a "shuttle":

$$\begin{gathered}
\text{symmetric polynomial of solutions of} \\
F_1(x_1, \dots, x_{K-1}) = 0, \dots, F_{K-1}(x_1, \dots, x_{K-1}) = 0 \\
\Downarrow \\
\mathcal{R}(F_1, \dots, F_{K-1}, F_K(x_1, \dots, x_{K-1})) \\
\Downarrow
\end{gathered}$$

$$\begin{gathered}\text{symmetric polynomial of solutions of}\\ \widetilde{F}_1(x_1,\dots,x_K)=0,\dots,\widetilde{F}_{K-1}(x_1,\dots,x_K)=0,\widetilde{F}_K(x_1,\dots,x_K)=0\\ \Downarrow\\ \mathcal{R}(\widetilde{F}_1,\dots,\widetilde{F}_{K-1},\widetilde{F}_K)\\ \Downarrow\\ \dots\end{gathered}$$

We will not treat the exceptional cases as when the system (2.10) or (2.18) has less than N solutions (some of the solutions become "infinite") or is incompatible, or has an infinite number of solutions (the eliminant in a particular variable may still exist). Neither will we discuss the case when the forms $F_{jn_j}(X)$ has a common zero with, say, $x_1 = 0$ and one has to find the eliminant in this variable: if one wants to find the eliminant in x for the system $F_1(x,y) := xy-1=0,\ F_2(x,y) := x^2y+x-2=0$ one has to take care of the case $x=0$ when the degrees of both polynomials decrease while the eliminants give rise to the "extraneous" solution $(0,0)$.

3. The Case of One Variable

As a basis for our approach to the optimization problem, we will first review some classical results due to Jacobi, Hermite, Sylvester and Kronecker on the relative distribution of the roots of two univariate polynomials in terms of the appropriate Hankel matrices (Krein and Naimark, 1981; Uteshev and Shulyak, 1992). For any real symmetric matrix M : $n_+(M)$ (or $n_-(M)$) denotes its positive (or negative) index, $r(M)$ its rank ($r(M) = n_+(M)+n_-(M)$), $\sigma(M)$ its signature ($\sigma(M) = n_+(M)-n_-(M)$), M_j its leading principal minor of the order j.

For the real polynomials (2.1) consider the following Laurent expansions

$$\frac{G(x)}{F(x)} = L(x) + \sum_{k=0}^{\infty}\frac{c_k}{x^{k+1}}, \tag{3.1}$$

$$\frac{F'(x)}{F(x)} = \sum_{k=0}^{\infty}\frac{s_k}{x^{k+1}}, \quad \frac{G(x)F'(x)}{F(x)} = L_1(x) + \sum_{k=0}^{\infty}\frac{h_k}{x^{k+1}}. \tag{3.2}$$

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

$$\frac{1}{F(x)} = \sum_{k=n-1}^{\infty}\frac{d_k}{x^{k+1}} \tag{3.3}$$

with the coefficients determined by the recurrent formulae

$$d_k = \begin{cases} 1/A_0 & \text{if } k = n-1\ ;\\ -(d_{n-1}A_1)/A_0 & \text{if } k = n\ ;\\ -(d_{k-1}A_1 + d_{k-2}A_2 + \dots + d_{n-1}A_{k-n+1})/A_0 & \text{if } k < 2n-1\ ;\\ -(d_{k-1}A_1 + d_{k-2}A_2 + \dots + d_{k-n}A_n)/A_0 & \text{if } k \ge 2n-1 \end{cases}$$

(we also set $d_\ell := 0$ for $\ell < n-1$) and then to multiply (3.3) by the corresponding polynomials. In this way, one can find the expressions for s_k, c_k and h_k which will be

rational functions with respect to the coefficients of F and G. On the other hand, these coefficients turn out to be symmetric functions of the roots $\lambda_1, \dots, \lambda_n$ of $F(x)$. So, for example,

$$c_k = \sum_{j=1}^{n} \frac{\lambda_j^k G(\lambda_j)}{F'(\lambda_j)} = \begin{cases} (d_{k+m}B_0 + d_{k+m-1}B_1 + \cdots + d_{n-1}B_{k+m-n+1}) & \text{if } k < n-1\ ; \\ (d_{k+m}B_0 + d_{k+m-1}B_1 + \cdots + d_k B_m) & \text{if } k \geq n-1 \end{cases} \tag{3.4}$$

with the first equality valid if all the roots λ_j are distinct. The coefficients of the expansions (3.2) can be obtained from c_k for particular choices of the polynomial G. In this way one can deduce the formulae (2.2) for the Newton sums s_k, while

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

Consider the following $n \times n$ Hankel matrices

$$S = [s_{j+k}]_{j,k=0}^{n-1}, \quad C = [c_{j+k}]_{j,k=0}^{n-1}, \quad H = [h_{j+k}]_{j,k=0}^{n-1} \tag{3.5}$$

and compute their leading principal minors $S_k, C_k, H_k,\ (k = 1, \dots, n)$.

Theorem 3.1. (Jacobi, Hermite, Sylvester) (1) *The number of distinct roots of $F(x)$ equals $r(S)$, and the number of distinct real roots of $F(x)$ equals $\sigma(S)$.*

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

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

We shall denote the last number by $\operatorname{nrr}\{F = 0 \mid G > 0\}$. For any real symmetric matrix M its positive and negative indices can be found with the help of the leading principal minors $M_1, \dots, M_r$:

$$n_+(M) = P(1, M_1, \dots, M_r), \quad n_-(M) = V(1, M_1, \dots, M_r) \quad (r := r(M)). \tag{3.6}$$

Here P (or V) is the number of permanences (or variations) of sign, and it was assumed that none of these minors vanishes. It is possible to generalize this rule to the case when $M_r \neq 0$, and there are no three consecutive zeros in this sequence. In particular, while computing using formulae (3.6), one may omit the minor equal to zero, provided that the neighboring ones do not vanish. If, in addition, the matrix M is of the Hankel type, then it is possible to extend this result to the case when any number of zeros exist in this sequence. For Frobenius' rule we refer to Gantmacher (1959). In the following, we will always assume, for simplicity, that for any matrix M under consideration there are no two consecutive zeros in the sequence $M_1, \dots, M_r$.

Corollary 3.1. *If all the numbers S_j, H_k do not vanish, then*

$$\operatorname{nrr}\{F = 0\} = P(1, S_1, \dots, S_n) - V(1, S_1, \dots, S_n), \tag{3.7}$$

$$\operatorname{nrr}\{F = 0 \mid G > 0\} = P(1, H_1, \dots, H_n) - V(1, S_1, \dots, S_n). \tag{3.8}$$

Theorem 3.2. *One has*

$$S_n := \det S = \prod_{1 \leq j < k \leq n} (\lambda_k - \lambda_j)^2 = \mathcal{D}(F)/A_0^{2n-2}, \tag{3.9}$$

$$H_n := \det H = S_n \prod_{1\le j\le n} G(\lambda_j) = \mathcal{R}(F,G) S_n / A_0^m. \tag{3.10}$$

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

THEOREM 3.3. (KRONECKER, 1897) *One has:*

$$\deg(\gcd(F,G)) = d \iff C_n = 0, \dots, C_{n-d+1} = 0, C_{n-d} \neq 0. \tag{3.11}$$

In this case, $\gcd(F,G)$ *equals the determinant obtained on replacing the last row in C_{n-d}:*

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

where $F_k(x) := A_0 x^{n-k} + A_1 x^{n-k-1} + \dots + A_{n-k}$. When $n > m$, the 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 $p_k(x) := c_0 x^k + \dots + c_k$.

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

$$p(x) := \varrho_0 x^{n-2} + \varrho_1 x^{n-3} + \dots + \varrho_{n-2}, \quad q(x) := q_0 x^{n-1} + q_1 x^{n-2} + \dots + q_{n-1}$$

satisfying the equality $p(x)F(x) + q(x)G(x) \equiv a_0 C_n$. On dividing the latter by $F(x)$ and expanding its both sides in series we obtain

$$p(x) + q(x)\left(\frac{c_0}{x} + \frac{c_1}{x^2} + \dots\right) \equiv a_0 C_n \left(\frac{d_{n-1}}{x^n} + \frac{d_n}{x^{n+1}} + \dots\right).$$

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

$$\begin{aligned} x^{-1} &: c_0 q_{n-1} + c_1 q_{n-2} + \dots + c_{n-1} q_0 = 0, \\ &\dots \\ x^{-n+1} &: c_{n-2} q_{n-1} + c_{n-1} q_{n-2} + \dots + c_{2n-3} q_0 = 0, \\ x^{-n} &: c_{n-1} q_{n-1} + c_n q_{n-2} + \dots + c_{2n-2} q_0 = C_n. \end{aligned}$$

Solving this system by Cramer's rule, we get the expressions for $q_0, \dots, q_{n-1}$, which coincide with those mentioned in Theorem 3.3. Similarly, one has

$$\begin{aligned} x^{n-2} &: \varrho_0 + q_0 c_0 = 0, \\ x^{n-3} &: \varrho_1 + q_0 c_1 + q_1 c_0 = 0, \end{aligned}$$

$$\begin{aligned} &\dots \\ 1 \quad &: \varrho_{n-2} + q_0 c_{n-2} + q_1 c_{n-3} + \dots + q_{n-2} c_0 = 0. \end{aligned} \tag{3.12}$$

From the above equations, the coefficients $\varrho_0, \dots, \varrho_{n-2}$ may be expressed in the determinantal form with the help of the representations of $q_0, \dots, q_{n-1}$ obtained earlier. However, for numerical computations of $\varrho_0, \dots, \varrho_{n-1}$ it would be easier to use formulae (3.12) directly.

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

Corollary 3.2. *One has the following relationship between the leading principal minors of the matrix C (or S) and subresultants (or subdiscriminants):*

$$A_0^{n+m} C_n = \mathcal{R}_0(F,G) = (-1)^{n(n-1)/2} \mathcal{R}(F,G), \quad A_0^{n+m-2k} C_{n-k} = \mathcal{R}_k(F,G) \quad (k > 0)$$

$$A_0^{2n-2k-2} S_{n-k} = n^{n-k-2} \mathcal{D}_k(F) \quad (0 < k < n-1). \tag{3.13}$$

By setting $d = 0$ in Theorem 3.3 *one can also find the linear representation of $\mathcal{R}(F,G)$:*

$$\mathcal{R}(F,G) \equiv (-1)^{n(n-1)/2} A_0^{n+m-1} (p(x)F(x) + q(x)G(x)) \tag{3.14}$$

provided that $m < n$. When $m \geq n$, 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 with the help of (3.4) *for the negative indices k).*

Corollary 3.3. *For the case $d = 1$, the single common root of $F(x)$ and $G(x)$ can be represented as*

$$x = s_1 - \frac{\Theta}{C_{n-1}}, \quad \text{where} \quad \Theta := \begin{vmatrix} c_0 & c_1 & \dots & c_{n-2} \\ c_1 & c_2 & \dots & c_{n-1} \\ \dots & & & \\ c_{n-3} & c_{n-2} & \dots & c_{2n-5} \\ c_{n-1} & c_n & \dots & c_{2n-3} \end{vmatrix}. \tag{3.15}$$

Moreover, since expansion (3.2) *is a particular case of* (3.1)*, one may also use in* (3.15) *the corresponding minors of the matrix H provided that $S_n \neq 0$ (see Theorem* 3.2*).*

Though this result is easily deduced from Theorem 3.3, an independent proof will be given (due to the present authors) which will be carried over to the bivariate case in Section 4. Let us assume, for simplicity, that all the roots $\lambda_1, \dots, \lambda_n$ of $F(x)$ are distinct, and let λ_1 be the root common with $G(x)$. Construct the polynomial $G_t(x) := (x-t)G(x)$. It is evident that for $t = \lambda_2, \dots, t = \lambda_n$ one has: $\deg(\gcd(F, G_t)) = 2$. According to (3.11), for the corresponding Hankel matrix

$$\widetilde{C}(t) := [c_{j+k+1} - c_{j+k} t]_{j,k=0}^{n-1}$$

the leading minors $\widetilde{C}_n(t)$ and $\widetilde{C}_{n-1}(t)$ vanish for those values of t.

One has: $\widetilde{C}_n(t) \equiv C_n F(t)$ and it is identically zero since $C_n = 0$. As a result of

elementary transformations, the minor $\widetilde{C}_{n-1}(t)$ can be represented in the form

$$\widetilde{C}_{n-1}(t) = (-1)^{n-1}\begin{vmatrix} c_0 & c_1 & \dots & c_{n-2} & 1 \\ c_1 & c_2 & \dots & c_{n-1} & t \\ \dots & & & & \dots \\ c_{n-2} & c_{n-1} & \dots & c_{2n-4} & t^{n-2} \\ c_{n-1} & c_n & \dots & c_{2n-3} & t^{n-1} \end{vmatrix}_{n\times n}$$
$$= (-1)^{n-1}(C_{n-1}t^{n-1} - \Theta t^{n-2} + \cdots).$$

Being a polynomial of the degree $n-1$, $\widetilde{C}_{n-1}(t)$ has $\lambda_2, \dots, \lambda_n$ as its roots. Formula (3.15) follows then from the two equalities:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = -A_1/A_0 = s_1, \quad \lambda_2 + \cdots + \lambda_n = \Theta/C_{n-1}.\square$$

We conclude the "classical" part of the present section with the following statement: Theorems 3.1 and 3.3 allow one to solve the problem of establishing the relative distribution of the roots of two polynomials that is usually treated in terms of euclidean algorithm, generalized polynomial remainder sequences or subresultants (Collins, 1971; Akritas, 1989; González-Vega *et al.*, 1990).

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

$$f(x) := a_0x^{n+1} + a_1x^n + \cdots + a_{n+1}, \quad a_0 < 0.$$

Then $\max f$ is attained in one of the critical points, i.e. on the real roots of $f'(x)$. If we denote the roots of $f'(x)$ by $\lambda_1, \dots, \lambda_n$, then $\max f$ is among the numbers $f(\lambda_1), \dots, f(\lambda_n)$. Let us find a polynomial $\mathcal{F}(z)$ having those numbers as its roots. That can be done purely algebraically: according to the results from Section 2, the polynomial

$$\mathcal{F}(z) := \mathcal{D}(f(x) - z) \tag{3.16}$$

(here the discriminant is evaluated with respect to the variable x) possesses the required property:

$$\mathcal{F}(z) \equiv (-1)^{(n-1)/2}(n+1)^{n+1}a_0^n(z - f(\lambda_1)) \dots (z - f(\lambda_n)).$$

By construction, its coefficients will be polynomial in $a_0, \dots, a_{n+1}$ with the constant term equal to $\mathcal{D}(f)$. If the symbolic expression for $\mathcal{D}(f)$ is known in terms of the coefficients of $f(x)$:

$$\mathcal{D}(f) := D(a_0, \dots, a_{n+1}),$$

then a compact formula for $\mathcal{F}(z)$ can be deduced via expanding $D(a_0, \dots, a_{n+1} - z)$ in powers of $(z - a_{n+1})$:

EXAMPLE 3.1. *For $f(x) = x^4 + a_2x^2 + a_3x + a_4$, one has*

$$\mathcal{D}(f) = -4a_2^3a_3^2 - 27a_3^4 + 16a_2^4a_4 - 128a_2^2a_4^2 + 144a_2a_3^2a_4 + 256a_4^3$$

and

$$\mathcal{F}(z) = -256(z-a_4)^3 - 128a_2^2(z-a_4)^2 - 16a_2(a_2^3 + 9a_3^2)(z-a_4) - a_3^2(4a_2^3 + 27a_3^2)$$
$$= -256z^3 + 128(6a_4 - a_2^2)z^2 + 16(16a_2^2a_4 - 48a_4^2 - a_2^4 - 9a_2a_3^2)z + \mathcal{D}(f).$$

However, for the general case, to obtain such a symbolic expression for $\mathcal{D}(f)$ is rather complicated, and thus, this approach, while possible in theory, becomes impractical. Keeping in mind that it is required not only to compute $\mathcal{F}(z)$, but to separate its roots, we state the following

PROBLEM 3.1. *For every real z find*

$$\operatorname{nrr}\{f'(x) = 0 \mid f(x) > z\}.$$

Let us use Theorems 3.1 and 3.3. Take $F(x) := f'(x), G(x) := f(x) - z$ and construct the matrix H as in (3.5). Consider the sequence of its leading principal minors

$$H_1(z), \ldots, H_n(z). \tag{3.17}$$

Let us investigate the features of this sequence. First of all, the leading coefficient of the polynomial $H_k(z)$ equals $(-1)^k S_k$, where S_k is the kth leading principal minor of the matrix S constructed as in (3.5) for $F(x) := f'(x)$. So, by formula (3.7), one is able to find $\operatorname{nrr}\{f'(x) = 0\}$.

ASSUMPTION 3.1. *Let $S_n \neq 0$.*

Thus, by Corollary 3.2, all the roots $\lambda_1, \ldots, \lambda_n$ of $f'(x)$ are distinct. By (3.10), one has $H_n(z) = S_n \prod_{j=1}^{n} (f(\lambda_j) - z)$, and thus the polynomial $H_n(z)$ differs from $\mathcal{F}(z)$ given by (3.16) only by the constant factor.

ASSUMPTION 3.2. *Let $\mathcal{D}(\mathcal{F}(z)) \neq 0$.*

Then all the critical values of $f(x)$ are distinct, and, by formula (3.8), we have the following result:

THEOREM 3.4. *Under Assumptions* 3.1 *and* 3.2*, one has*

$$\operatorname{nrr}\{\mathcal{F}(z) = 0\} = \operatorname{nrr}\{f'(x) = 0\}; \tag{3.18}$$

$$\begin{aligned} \operatorname{nrr}\{\mathcal{F}(z) = 0 \mid a < z < b\} &= \operatorname{nrr}\{f'(x) = 0 \mid a < f(x) < b\} \\ &= P(1, H_1(a), \ldots, H_n(a)) - P(1, H_1(b), \ldots, H_n(b)). \end{aligned} \tag{3.19}$$

This theorem implies that sequence (3.17) is the Sturm's series for $\mathcal{F}(z)$, by means of which the critical values of $f(x)$ can 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 can also be found via the minors of the matrix $H(z)$. Indeed, by Corollary 3.3, x can be expressed by the formula (3.15), in which one should set $s_1 := -na_1/[(n+1)a_0]$, $c_k := h_k(z)$, and $C_{n-1} := H_{n-1}(z)$.

What happens if Assumption 3.2 is violated? The condition $\mathcal{D}(\mathcal{F}(z)) = 0$ is equivalent to the existence of two equal critical values for $f(x)$. One cannot claim any longer that the maximal real root of $\mathcal{F}(z)$ coincides with $\max f(x)$:

EXAMPLE 3.2. *Find the* $\max f$ *for* $f(x) = -x^6 - 135x^2 - 324x$.

Here the polynomial

$$\mathcal{F}(z) = -46\,656\,(z - 540)^2(z^3 + 1080z^2 + 1603\,800z - 354\,294\,000)$$

has the maximal root $z = 540$; however, it corresponds to the *imaginary* roots $\lambda_{1,2} = (-3 \pm i\sqrt{15})/2$ of the derivative $f'(x) = -6(x^5 + 45x + 54)$. As a matter of fact, $\max f(x) = 90(-4 + 5\sqrt[3]{10} - \sqrt[3]{100}) = 191.75261 \pm 10^{-5}$ is attained at the root $\lambda_3 = 1 - \sqrt[3]{10} = -1.15443 \pm 10^{-5}$.

So, for this example formulae (3.18) and (3.19) will not work, since polynomials $f'(x)$ and $\mathcal{F}(z)$ have different numbers of real roots. Nevertheless, it would be interesting to take a look at the behaviour of sequence (3.17):

$$H_1(z) = -5z;\ H_2(z) = -81000z;\ H_3(z) = 7290000(-2z^2 + 2835z - 583\,200);$$
$$H_4(z) = 14\,580\,000(z - 540)(2z^3 + 1485z^2 + 2\,114\,100z - 1766\,549\,250);$$
$$H_5(z) = 73\,811\,250\,000/46\,656\mathcal{F}(z).$$

The difference (3.19) computed for any interval $]a, b[$, where $0 < a < b < +\infty, a \neq 540, b \neq 540$ locates the single critical value for $f(x)$ lying in $]191, 192[$ and ignores $z = 540$. Thus, sequence (3.17) is intelligent enough to neglect the roots of $\mathcal{F}(z)$ corresponding to the imaginary roots of $f'(x)$.□

On the other hand, this sequence reacts properly when $f(x)$ assumes the same critical value in two (or more) *real* critical points:

EXAMPLE 3.3. *Find the* $\max f$ *for* $f(x) = -x^6 - 10x^3 + 12x$.

For this example formula (3.19) applied to the sequence

$$H_1(z) = -5(z - 15); H_2(z) = 1000(z - 15);$$
$$H_3(z) = 125(9z^3 - 585z^2 + 13175z - 32375);$$
$$H_4(z) = -1250(z - 5)(27z^3 - 1930z^2 + 44775z - 29000);$$
$$H_5(z) = 625000(z - 5)^2(z^3 - 65z^2 + 1200z + 8000)$$

permits one to find out that for any interval $a < z < b$ with $0 < a < 5, b > 5$ there correspond two real critical points. Using Theorem 3.3, one may obtain a quadratic equation for their evaluation: $\max f = 5$ is attained at $\lambda_{1,2} = (-1 \pm \sqrt{5})/2$. □

It is desirable to check Assumption 3.2 in terms of the matrices S and H (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 3.5. $\mathcal{D}(\mathcal{F}(z)) = k[\mathcal{D}(f')]^3\Phi^2(A_0, \ldots, A_n)$, *where* k *is a constant depending only on* n, *and* Φ *is a real form of the degree* $3(n-1)(n-2)/2$ *with respect to the coefficients of* $f'(x)$.

PROOF. If $f(x) = a_0x^{n+1} + \cdots + a_{n+1}$, $f'(x) = A_0x^n + \cdots + A_n$ and $f'(\lambda_j) = f'(\lambda_k) = 0$, then $f(\lambda_k) - f(\lambda_j) = (\lambda_k - \lambda_j)^3\Psi(\lambda_j, \lambda_k)$,where

$$\Psi(\lambda_j, \lambda_k) := 1/2\big[a_{n-2} + a_{n-3}(2\lambda_j + 2\lambda_k) + a_{n-4}(3\lambda_j^2 + 4\lambda_j\lambda_k + 3\lambda_k^2)$$
$$+ a_{n-5}(4\lambda_j^3 + 6\lambda_j^2\lambda_k + 6\lambda_j\lambda_k^2 + 4\lambda_k^3) + \cdots\big].$$

(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 symmetric 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 $A_0, \dots, A_n$. Hence, we have

$$\begin{aligned}\mathcal{D}(\mathcal{F}(z)) &= ((-1)^{(n-1)/2}(n+1)A_0^n)^{2n-2} \prod_{1\le j<k\le n} (f(\lambda_k) - f(\lambda_j))^2 \\ &= (n+1)^{2(n-1)} \left(A_0^{2n-2} \prod_{1\le j<k\le n} (\lambda_k - \lambda_j)^2 \right)^3 \left(A_0^{(n-1)(n-3)} \prod_{1\le j<k\le n} \Psi(\lambda_j, \lambda_k) \right)^2 \\ &= k[\mathcal{D}(f'(x))]^3 \Phi^2(A_0, \dots, A_n).\end{aligned}$$

Here polynomial Φ is a form of the degree $3(n-1)(n-2)/2$ and of the weight $n(n-1) \times (n-2)/2$ with respect to $A_0, \dots, A_n$. Its coefficients may be made integers by an appropriate choice of the constant k. It turns out that

$\Phi \equiv 1$ for $n = 2$;
$\Phi \equiv 27A_0^2A_3 + 2A_1^3 - 9A_0A_1A_2 \equiv -\mathcal{R}(f', f''')/(8A_0)$ for $n = 3$;
$\Phi \equiv \mathcal{D}\big(5[f''']^2 - 6f'f^{(5)}\big)/A_0^{7-n}$ for $n = 4$ and for $n = 5$. □

4. The Case of Two Variables

As in Section 3, we shall begin with the "classical" part, i.e. a review of Hermite's method for the separation of solutions of a system of polynomial equations. We will keep the notation and assumptions of Section 2, but in the present section only real polynomials will be considered.

The general separation problem was formulated in Section 1 as that of finding the

$$\operatorname{nrs}\{(2.10) \mid G(x, y) > 0\},$$

i.e. the number of distinct real solutions of (2.10) that satisfy the inequality

$$G(x, y) := \sum_{\substack{p,\ell=0 \\ p+\ell\le m}} b_{p\ell} x^p y^\ell > 0. \tag{4.1}$$

To find this number, we need to evaluate symmetric functions of the solutions of the system (2.10). For this aim, we will use the method invented by Jacobi in 1835–36 (which differ from that mentioned in the proof of Theorem 2.1).

The first step is to find the eliminants $\mathcal{X}(x)$ and $\mathcal{Y}(y)$ defined by (2.12). To do this, one may now employ Corollary 3.2. 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, neglecting the sign, one gets

$$\mathcal{Y}(y) \equiv A_{1,0}^{n_1+n_2} \det[c_{j+k}(y)]_{j,k=0}^{n_1-1}. \tag{4.2}$$

Moreover, by means of Corollary 3.2, one can also find its linear representation, i.e. polynomials $\mathcal{P}(x, y)$ and $\mathcal{Q}(x, y)$ from $\mathbb{R}[x, y]$ such that

$$\mathcal{P}(x, y)F_1 + \mathcal{Q}(x, y)F_2 \equiv \mathcal{Y}(y). \tag{4.3}$$

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

$$\mathcal{M}(x, y)F_1 + \mathcal{N}(x, y)F_2 \equiv \mathcal{X}(x). \tag{4.4}$$

Sometimes equalities (4.3) and (4.4) are referred to as the *Bézout identities*. The degrees of the polynomials $\mathcal{M}, \mathcal{N}, \mathcal{P}$ and $\mathcal{Q}$ constructed according to Theorem 3.3 will satisfy the following restrictions (Uteshev and Shulyak, 1992):

$$\left\{\begin{array}{rclcl} \deg \mathcal{M} & = & \deg_x \mathcal{M} & \le & N-n_1,\ \deg_y \mathcal{M} \le n_2-1, \\ \deg \mathcal{N} & = & \deg_x \mathcal{N} & \le & N-n_2,\ \deg_y \mathcal{N} \le n_1-1, \\ \deg \mathcal{P} & = & \deg_y \mathcal{P} & \le & N-n_1,\ \deg_x \mathcal{P} \le n_2-1, \\ \deg \mathcal{Q} & = & \deg_y \mathcal{Q} & \le & N-n_2,\ \deg_x \mathcal{Q} \le n_1-1. \end{array}\right. \tag{4.5}$$

Introducing now an important function

$$\mathcal{V}(x,y) := \mathcal{M}(x,y)\mathcal{Q}(x,y) - \mathcal{N}(x,y)\mathcal{P}(x,y)$$

we can deduce from (4.5) the restrictions:

$$\deg \mathcal{V} \le 2N - n_1 - n_2, \quad \deg_x \mathcal{V} \le N-1, \quad \deg_y \mathcal{V} \le N-1. \tag{4.6}$$

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

$$\frac{\mathcal{V}(x,y)}{\mathcal{X}(x)\mathcal{Y}(y)} = \sum_{j,k=0}^{\infty} \frac{d_{jk}}{x^{j+1}y^{k+1}} \tag{4.7}$$

(using expansions of the type (3.3)). Because of the restrictions (4.6), we have:

$$d_{jk} = 0 \text{ for } j+k \le n_1+n_2-3. \tag{4.8}$$

On the other hand, it turns out that the expansion (4.7) plays a role similar to that of expansion (3.3). Indeed, let us multiply (4.7) first by the Jacobian

$$\mathcal{J}(x,y) := \partial F_1/\partial x \cdot \partial F_2/\partial y - \partial F_2/\partial x \cdot \partial F_1/\partial y, \tag{4.9}$$

then by $G(x,y)$, and, finally, by their product:

$$\frac{\mathcal{J}\cdot\mathcal{V}}{\mathcal{X}\cdot\mathcal{Y}} = L_0(x,y) + \sum_{j,k=0}^{\infty} \frac{s_{jk}}{x^{j+1}y^{k+1}}\ , \quad \frac{G\cdot\mathcal{V}}{\mathcal{X}\cdot\mathcal{Y}} = L_1(x,y) + \sum_{j,k=0}^{\infty} \frac{c_{jk}}{x^{j+1}y^{k+1}} \tag{4.10}$$

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

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

$$c_{jk} = \sum_{\substack{p,\ell=0 \\ p+\ell\le m}} b_{p\ell} d_{p+j,\ell+k}.$$

Formula for s_{jk} can be obtained from this one for $G := \mathcal{J}$, while

$$h_{jk} = \sum_{\substack{p,\ell=0 \\ p+\ell\le m}} b_{p\ell} s_{p+j,\ell+k}. \tag{4.11}$$

By construction, all these numbers will be rational functions with respect to the coefficients of the polynomials F_1, F_2 and G. On the other hand, they proved to be the values of appropriate symmetric functions of solutions of the system (2.10):

THEOREM 4.1. (JACOBI) *One has:*

$$s_{jk} = \sum_{p=1}^{N} \alpha_p^j \beta_p^k, \quad h_{jk} = \sum_{p=1}^{N} G(\alpha_p, \beta_p) \alpha_p^j \beta_p^k. \tag{4.12}$$

If $\mathcal{J}(\alpha_p, \beta_p) \neq 0$ *for all* $p = 1, \ldots, N$ *(i.e. all the solutions of* (2.10) *are simple), then*

$$c_{jk} = \sum_{p=1}^{N} \frac{\alpha_p^j \beta_p^k G(\alpha_p, \beta_p)}{\mathcal{J}(\alpha_p, \beta_p)}, \tag{4.13}$$

$$d_{jk} = \sum_{p=1}^{N} \frac{\alpha_p^j \beta_p^k}{\mathcal{J}(\alpha_p, \beta_p)} \tag{4.14}$$

and $= 0$ *for* $j + k \leq n_1 + n_2 - 3$ *(v.* (4.8)*).*

For the proof of the above result we refer to the Appendix of the paper by Uteshev and Shulyak (1992).

To find $\mathrm{nrs}\{(2.10) \mid G(x, y) > 0\}$ we shall consider two different approaches. The first one is based on the following

ASSUMPTION 4.1. *Let* $\mathcal{D}(\mathcal{X}(x)) \neq 0$.

Under this assumption, all the solutions (α_j, β_j) of (2.10) are distinct, they have distinct x-components, and

$$\mathrm{nrs}\{(2.10)\} = \mathrm{nrr}\{\mathcal{X}(x) = 0\}$$

where the latter number can be found by Theorem 3.1.

THEOREM 4.2. (BRIOSCHI) *Under Assumptions* 2.1, 2.2 *and* 4.1, *one has*

$$\mathrm{nrs}\{(2.10) \mid G(x, y) > 0\} = n_+(H) - (N - \mathrm{nrs}\{(2.10)\})/2, \tag{4.15}$$

where the $N \times N$ *Hankel matrix* H *is defined by*

$$H = [h_{j+k,0}]_{j,k=0}^{N-1}. \tag{4.16}$$

COROLLARY 4.1. *One has:*

$$H_N := \det H = \prod_{1 \leq j < k \leq N} (\alpha_k - \alpha_j)^2 \prod_{1 \leq p \leq N} G(\alpha_p, \beta_p) = \frac{\mathcal{D}(\mathcal{X}(x))}{\mathcal{A}_0^{2N-2}} \prod_{1 \leq p \leq N} G(\alpha_p, \beta_p).$$

In the original work by Hermite the idea of another approach can also be found (Uteshev and Cherkasov, 1996). As a matter of fact, it is possible to replace Assumption 4.1 by one, concerning only the leading forms (2.11) of the polynomials F_1 and F_2. We shall sketch this approach only for a particular case when $n_1 = n_2 := n$.

ASSUMPTION 4.2. *Let all the subresultants* $\mathcal{R}_1(\mathcal{A}_0), \ldots, \mathcal{R}_{n-1}(\mathcal{A}_0)$ *of the resultant (2.13) be nonzero.*

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

$$C := [C_{p\ell}]_{p,\ell=0}^{n-1}, \quad \text{where } C_{p\ell} := [c_{j+k,p+\ell}]_{j=\overline{0,2(n-p-1)};k=\overline{0,2(n-\ell-1)}} \tag{4.17}$$

$$H := [H_{p\ell}]_{p,\ell=0}^{n-1}, \quad \text{where } H_{p\ell} := [h_{j+k,p+\ell}]_{j=\overline{0,2(n-p-1)};k=\overline{0,2(n-\ell-1)}} \tag{4.18}$$

while S is constructed similarly from the coefficients of the corresponding expansion (4.10). This structure of the matrices will be explained in the proof of the following

THEOREM 4.3. *Under Assumptions* 2.1, 2.2 *and* 4.2, *one has:*

(1) *The number of distinct solutions of* (2.10) *equals* $r(S)$ *and* $\mathrm{nrs}\{(2.10)\} = \sigma(S)$.

(2) *The number of solutions of* (2.10) *satisfying the condition* $G(x,y) = 0$ *equals the deficiency index of the matrix* (4.17), i.e. $N - r(C)$. *If a solution of* (2.10) *of multiplicity* m_1 *that is, at the same time, a zero for* $G(x,y)$ *of multiplicity* m_2 *exists it should be counted* $\min(m_1, m_2)$ *times.*

(3) *If the matrix* H *constructed by* (4.18) *is nonsingular, then all the solutions of* (2.10) *are distinct and formula* (4.15) *remains valid.*

PROOF. We shall present here only an idea of the proof for part (3) and, simultaneously, for Theorem 4.2. For both cases, the starting point is the following quadratic form in the real variables $x_0, \ldots, x_{N-1}$:

$$\sum_{j=1}^{N} G(\alpha_j, \beta_j)[x_0\Psi_0(\alpha_j, \beta_j) + x_1\Psi_1(\alpha_j, \beta_j) + \cdots + x_{N-1}\Psi_{N-1}(\alpha_j, \beta_j)]^2 \tag{4.19}$$

where the system of monomials $\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1}$ is linearly independent on solutions of the system (2.10). Under Assumption 4.1, one may take $\Psi_k(\alpha,\beta) := \alpha^k$, and then the matrix of the form (4.19) coincides with (4.16). Under Assumption 4.2, one should take the following system

$$\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1} := \{\alpha^p\beta^q \mid 0 \le q \le n-1, 0 \le p \le 2n-2q-2\} \tag{4.20}$$
$$= \left\{ \begin{array}{ccccccccc} 1, & \alpha, & \alpha^2, & \ldots, & \ldots, & \alpha^{2n-4}, & \alpha^{2n-3}, & \alpha^{2n-2}, \\ \beta, & \alpha\beta, & \alpha^2\beta, & \ldots, & \alpha^{2n-3}\beta, & \alpha^{2n-4}\beta, & & \\ \beta^2, & \alpha\beta^2, & \ldots, & \alpha^{2n-6}\beta^2, & & & & \\ \ldots, & & & & & & & \\ \beta^{n-1} & & & & & & & \end{array} \right\}.$$

Let us prove that

$$(\det[\Psi_{k-1}(\alpha_j,\beta_j)]_{k,j=1}^{N})^2 = \det D \prod_{1\le p\le N} \mathcal{J}(\alpha_p,\beta_p) \tag{4.21}$$

where the $N \times N$ matrix D is made of the coefficients of the expansion (4.7) similarly to matrix (4.18). Suppose first, that $\mathcal{J}(\alpha_j,\beta_j) \neq 0$ for all $j = 1, \ldots, N$. The left-hand side of (4.21) can be represented as

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

where ^{t} denotes the transposition of the matrix. By the equality (4.14), one then has the validity of (4.21). Thus, it is true under the assumption when all the solutions of (2.10) are simple. The set of the systems (2.10) possessing a multiple solution is of measure zero in the space of the coefficients of F_1 and F_2. The equality (4.21) is an algebraic one with respect to those coefficients. Hence, it will also be true for the general case.

Due to the equality (4.8), matrix D may be put into the block triangular form by

means of some elementary transformations of its rows and columns. The blocks on the diagonal happen to be of the Hankel type

$$\det[d_{2n-j-k-2,j+k}]_{j,k=0}^{p-1} = \mathcal{R}_p(\mathcal{A}_0)/\mathcal{A}_0 \quad \text{for } p = 1, \dots, n-1,$$

$$\det[d_{2n-j-k-2,j+k}]_{j,k=0}^{n-1} = 1/\mathcal{A}_0. \tag{4.22}$$

So, under Assumption 4.2, the determinant on the left-hand side of (4.21) does not vanish, and the monomial system (4.20) is linearly independent on the solutions of (2.10) .

Formula (4.15) then follows from the inertia law for quadratic forms.□

Corollary 4.2. *One has*

$$\det C = \Upsilon \prod_{1\le p\le N} G(\alpha_p, \beta_p), \tag{4.23}$$

$$\det S = \Upsilon \prod_{1\le p\le N} \mathcal{J}(\alpha_p, \beta_p), \quad \det H = \det S \prod_{1\le p\le N} G(\alpha_p, \beta_p), \tag{4.24}$$

where $\Upsilon = (-1)^{(n-1)(n-2)/2}[\mathcal{R}_1(\mathcal{A}_0)\dots\mathcal{R}_{n-1}(\mathcal{A}_0)]^2/\mathcal{A}_0^{2n-1}$.

Remark. As a matter of fact, the only essential assumption in the above theorem is Assumption 2.2, whereas Assumption 4.2 is not. The monomial basis $\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1}$ could have been composed in accordance with the position of nonzero Hankel minors of determinant (4.22). We chose basis (4.20) only because of its elegant relationship with subresultants. For the general case when $n_1 := \deg F_1 \le n_2 := \deg F_2$, it will take the form:

$$\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1} := \{\alpha^p\beta^q \mid 0 \le q \le n_1 - 1, 0 \le p \le n_1 + n_2 - 2q - 2\}$$

while in Hermite's original paper the basis

$$\{\Psi_k(\alpha,\beta)\}_{k=0}^{N-1} := \{\alpha^p\beta^q \mid 0 \le p \le n_1 - 1, 0 \le q \le n_2 - 1\} \tag{4.25}$$

was considered.

The second part of Theorem 4.3 gives us a condition for the existence of a common zero for $F_1(x,y), F_2(x,y)$ and $G(x,y)$. If that zero is unique, then it can be expressed as a rational function of the coefficients of the polynomials considered. For this aim, construct a new parameter-dependent matrix of the same structure as matrix (4.17):

$$\widetilde{C} := [\widetilde{C}_{p\ell}]_{p,\ell=0}^{n-1}, \quad \text{where } \widetilde{C}_{p\ell} := [c_{j+k+1,p+\ell} - tc_{j+k,p+\ell}]_{j=\overline{0,2(n-p-1)};k=\overline{0,2(n-\ell-1)}}. \tag{4.26}$$

So, the coefficients of t are the entries of the matrix $(-C)$ with C defined by (4.17). Expand the $(N-1)$th leading principal minor of (4.26) into powers of t:

$$\widetilde{C}_{N-1}(t) = (-1)^{N-1}(C_{N-1}t^{N-1} + \Theta t^{N-2} + \cdots).$$

Theorem 4.4. *Let Assumptions* 2.1, 2.2 *and* 4.2 *be fulfilled, and* $C_N = 0, C_{N-1} \ne 0$. *Then the* x*-component of a single common zero for* $F_1(x,y), F_2(x,y)$ *and* $G(x,y)$ *can be found by the formula:*

$$x = s_{1,0} + \Theta/C_{N-1}. \tag{4.27}$$

Proof. The idea is the same as in the proof of Corollary 3.3. Let $\Lambda_1 = (\alpha_1, \beta_1)$ be the

single common zero for F_1, F_2, G, and assume, for simplicity, that all the solutions for the system (2.10) have distinct x-components. Construct a parameter-dependent polynomial $G_t(x,y) := (x-t)G(x,y)$. It has two zeros common with $F_1(x,y)$ and $F_2(x,y)$ when the parameter t takes the values $\alpha_2, \ldots, \alpha_N$. The matrix C constructed for F_1, F_2 and G_t coincides now with $\widetilde{C}$ introduced by (4.26). By Theorem 4.3, the two greatest leading principal minors of that matrix must vanish for those $N-1$ values of t. Being a polynomial of the degree $N-1$, $\widetilde{C}_{N-1}(t)$ has $\alpha_2, \ldots, \alpha_N$ as its roots. Formula (4.27) follows then from the two equalities:

$$\alpha_1 + \cdots + \alpha_N = s_{1,0}, \quad \alpha_2 + \cdots + \alpha_N = -\Theta/C_{N-1}. \qquad \square$$

Remark. To obtain the y-component for the common zero one should construct the matrix $\widetilde{C}$ for $G_u(x,y) := (y-u)G(x,y)$. Let us combine both procedures and construct the matrix $\widetilde{C}$ for $G_{t,u}(x,y) := (x-t)(y-u)G(x,y)$. Then the expansion of its $(N-1)$-th leading principal minor in powers of t and u

$$\widetilde{C}_{N-1}(t,u) = C_{N-1}t^{N-1}u^{N-1} + \Theta t^{N-2}u^{N-1} + \Xi t^{N-1}u^{N-2} + \cdots$$

permits one to find both components of a common zero simultaneously:

$$x = s_{1,0} + \Theta/C_{N-1}, \quad y = s_{0,1} + \Xi/C_{N-1}. \tag{4.28}$$

Under the additional assumption that $S_N \neq 0$, one may also use in (4.28) the corresponding minors of the matrix H.

Let us now apply the above results to 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 tackle 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 leading form

$$f_{(n+1)}(x,y) := a_{n+1,0}x^{n+1} + a_{n,1}x^n y + \cdots + a_{0,n+1}y^{n+1} \tag{4.29}$$

of an even degree $n+1$. This form also has to satisfy the following

Assumption 4.3. *Let the polynomial $f_{(n+1)}(1,y)$ be negative for every $y \in \mathbb{R}$.*

The latter condition holds iff $a_{0,n+1} < 0$ and $f_{(n+1)}(1,y)$ does not have real roots; thus, it can be verified by Theorem 3.1.

Consider the system

$$\partial f/\partial x = 0, \quad \partial f/\partial y = 0 \tag{4.30}$$

yielding the critical points of $f(x,y)$.

As in the previous section, let us state the following

Problem 4.1. *For every real z find*

$$\mathrm{nrs}\{(4.30) \mid f(x,y) > z\} \tag{4.31}$$

In order to apply the preceding results to solving this problem, we have to imply

some restrictions on $f(x,y)$ guaranteeing the fulfilment of Assumption 2.2 and either of Assumptions 4.1 or 4.2 for the system (4.30). We will restrict ourselves to the approach based on Theorem 4.3.

ASSUMPTION 4.4. *Let the leading form* (4.29) *satisfy the following conditions: the polynomial* $\varphi(y) := f_{(n+1)}(1,y)$ *has nonzero subdiscriminants and*

$$(-1)^{(n+1)/2}\mathcal{D}(\varphi) > 0, \quad P\big(1, n+1, \mathcal{D}_{n-1}(\varphi), \dots, \mathcal{D}_1(\varphi), \mathcal{D}(\varphi)\big) = (n+1)/2. \tag{4.32}$$

Under these conditions, Assumptions 2.2 and 4.2 are satisfied because of the relationship between the (sub)resultant and (sub)discriminant (2.8). Thus, for system (4.30), formulae (2.13) can be rewritten in the form

$$N := n^2, \quad \mathcal{A}_0 := (-1)^{(n+1)/2}(n+1)^{n-1}\mathcal{D}(\varphi(y)). \tag{4.33}$$

Furthermore, Assumption 4.3 is also satisfied because of formulae (3.7), (3.13) and the equality from (4.32).

By now applying Theorem 4.3 for the case $F_1 := \partial f/\partial x, F_2 := \partial f/\partial y$, and $G := f(x,y) - z$, one can solve Problem 4.1 in exactly the same manner as Problem 3.1. Calculating the leading principal minors for matrix (4.18) we obtain the polynomial sequence in z:

$$H_1(z), \dots, H_N(z). \tag{4.34}$$

The leading coefficient of the polynomial $H_k(z)$ equals $(-1)^k S_k$, where S_k is the kth leading principal minor of the matrix S constructed similarly to (4.17) from the coefficients s_{jk} of the expansion (4.10). The Jacobian (4.9) now coincides with the Hessian of $f(x,y)$:

$$\mathcal{H}(f) := (\partial^2 f/\partial x^2)(\partial^2 f/\partial y^2) - (\partial^2 f/[\partial x \partial y])^2. \tag{4.35}$$

So, by part (1) of Theorem 4.3, one can find $\mathrm{nrs}\{(4.30)\}$.

ASSUMPTION 4.5. *Let* $S_N \neq 0$.

Under this condition, all the critical points of $f(x,y)$ are simple. By (4.24), the polynomial

$$\mathcal{F}(z) := (-1)^N H_N(z)/S_N \equiv (z - f(\alpha_1, \beta_1)) \dots (z - f(\alpha_N, \beta_N)) \tag{4.36}$$

has roots coinciding with the critical values of f.

ASSUMPTION 4.6. *Let* $\mathcal{D}(\mathcal{F}(z)) \neq 0$.

Under this assumption, there exists a one-to-one correspondence between the real solutions of (4.30) and the real roots of $\mathcal{F}(z)$. Using part (2) of Theorem 4.3, we get the following result:

THEOREM 4.5. *Under Assumptions 4.4–4.6, one has*

$$\begin{aligned} &\mathrm{nrr}\{\mathcal{F}(z) = 0 \mid a < z < b\} = \mathrm{nrs}\{(4.30) \mid a < f(x,y) < b\} \\ &= P(1, H_1(a), \dots, H_N(a)) - P(1, H_1(b), \dots, H_N(b)). \end{aligned} \tag{4.37}$$

Thus, sequence (4.34) turns out to be the Sturm series for $\mathcal{F}(z)$, i.e. it permits to localize the real roots of that polynomial. Once the greatest (real) root of $\mathcal{F}(z)$ is found, one can get the coordinates of corresponding critical point (α, β) of $f(x, y)$ with the help of the remark after Theorem 4.4.

EXAMPLE 4.1. *Find the* $\max f$ *for*

$$\begin{aligned} f(x,y) := &- x^4 + 8x^3y - 24x^2y^2 + 24xy^3 - 8y^4 \\ &- 4/3x^3 - 8x^2y + 24xy^2 - 56/3y^3 + 10x^2 + 16xy + 60x + 32y. \end{aligned}$$

The leading form $f_{(4)}(x, y)$ satisfies Assumption 4.4, thus we may use Theorem 4.3. To find the entries of the matrix (4.18), we first evaluate s_{jk} for the system (4.30)

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

and then use formula (4.11). The sequence (4.34):

$$\begin{aligned}
H_1(z) &= -1125\,z + 662\,821, \\
H_2(z) &= 2136\,375\,z^2 - 4092\,726\,186\,z + 870894052211, \\
H_3(z) &= -57\,141\,430\,875\,z^3 + \cdots, \\
H_4(z) &= 1227\,163\,832\,960\,625\,z^4 + \cdots, \\
H_5(z) &= -14\,616\,183\,736\,762\,689\,375\,z^5 + \cdots, \\
H_6(z) &= 38\,363\,368\,500\,598\,188\,375\,z^6 + \cdots, \\
H_7(z) &= -72\,729\,653\,454\,356\,625\,z^7 + \cdots, \\
H_8(z) &= \frac{79\,164\,837\,199\,872}{48\,828\,125}\,(1250198701723125\,z^8 - 6364\,358\,382\,211\,742\,610\,z^7 \\
&\quad +9351\,549\,963\,140\,311\,543\,266\,z^6 \\
&\quad -3802\,534\,247\,698\,983\,423\,134\,442\,z^5 + 183\,480\,845\,901\,538\,243\,869\,874\,764\,z^4 \\
&\quad +89\,292\,706\,735\,989\,389\,296\,738\,993\,578\,z^3 \\
&\quad +1048\,657\,283\,190\,908\,842\,923\,598\,785\,102\,z^2 \\
&\quad -294\,298\,830\,961\,184\,186\,968\,080\,427\,432\,494\,z \\
&\quad -7351\,442\,540\,949\,758\,064\,538\,454\,022\,899\,057), \\
H_9(z) &= -2460\,375\,z^9 + 13\,046\,305\,743\,z^8 - 20\,953\,332\,885\,564\,z^7 \\
&\quad +10\,858\,379\,628\,617\,100\,z^6 \\
&\quad -1199\,221\,437\,495\,632\,850\,z^5 - 369\,773\,782\,407\,882\,562\,734\,z^4 \\
&\quad +33\,574\,934\,405\,487\,787\,787\,124\,z^3 \\
&\quad +8363\,310\,121\,361\,184\,850\,064\,700\,z^2 + 438\,702\,308\,762\,940\,646\,094\,396\,529\,z \\
&\quad +6672\,685\,776\,490\,188\,470\,056\,561\,943.
\end{aligned}$$

All the minors $H_k(z)$, except for $H_8(z)$, were divided by positive rationals, and, up to those numbers, their leading coefficients coincide with the leading principal minors S_k for the matrix S from Theorem 4.3. According to the first statement of this theorem, all the nine critical points of $f(x, y)$ are real and distinct. Using formulae (4.37), one can separate the roots of $H_9(z)$ including the maximal one:

$$\max f = 2797.86763 \pm 10^{-5}.$$

To evaluate the corresponding critical point (α, β), construct the matrix $\widetilde{H}(z,t,u)$ having the same structure as (4.18) with the elements

$$\widetilde{h}_{k\ell}(z) := tu\ h_{k\ell}(z) - t\ h_{k,\ell+1}(z) - u\ h_{k+1,\ell}(z).$$

Expand its leading principal minor $\widetilde{H}_8$ in decreasing powers of t and u:

$$\widetilde{H}_8(z,t,u) = H_8(z)t^8u^8 + \Theta(z)t^7u^8 + \Xi(z)t^8u^7 + \cdots$$

where

$$\begin{aligned}
\Theta(z) = \frac{29\,686\,813\,949\,952}{244\,140\,625}\,(&298\,578\,140\,055\,275\,625\,z^8\\
&-1509\,754\,934\,694\,047\,566\,680\,z^7\\
&+2182\,041\,088\,055\,721\,119\,565\,036\,z^6\\
&-834\,972\,939\,779\,438\,151\,129\,392\,712\,z^5\\
&+17\,334\,377\,438\,378\,913\,156\,914\,869\,734\,z^4\\
&+21\,171\,009\,495\,975\,011\,130\,288\,331\,602\,648\,z^3\\
&+969\,436\,681\,430\,804\,531\,145\,596\,694\,774\,732\,z^2\\
&-16\,802\,891\,776\,098\,857\,470\,342\,200\,724\,414\,904\,z\\
&-804\,885\,536\,988\,530\,448\,655\,114\,307\,218\,550\,407),\\
\Xi(z) = \frac{4947\,802\,324\,992}{244\,140\,625}\,(&1487\,202\,210\,104\,653\,125\,z^8\\
&-7529\,701\,417\,689\,854\,226\,600\,z^7\\
&+10\,917\,360\,698\,419\,349\,845\,101\,324\,z^6\\
&-4228\,408\,213\,906\,151\,771\,787\,103\,128\,z^5\\
&+111\,495\,525\,629\,039\,164\,175\,673\,846\,366\,z^4\\
&+105\,589\,619\,054\,271\,085\,959\,982\,676\,340\,072\,z^3\\
&+4146\,820\,611\,907\,993\,982\,121\,191\,327\,494\,668\,z^2\\
&-134\,193\,165\,831\,965\,927\,253\,094\,323\,259\,454\,376\,z\\
&-4898\,904\,812\,405\,735\,227\,047\,915\,809\,040\,587\,323).
\end{aligned}$$

If z is a root of $H_9(z)$, then the corresponding (α, β) can be found by formulae (4.28)

$$\alpha = s_{1,0} + \Theta(z)/H_8(z)\ , \quad \beta = s_{0,1} + \Xi(z)/H_8(z). \tag{4.38}$$

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

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

Remark 4.1. In order to simplify the calculations, one may replace the polynomial condition $f(x,y) > z$ by one of lower degree:

$$\mathrm{nrs}\{(4.30) \mid f(x,y) > z\} = \mathrm{nrs}\{(4.30) \mid f_*(x,y) := f - \tfrac{1}{n+1}(x\partial f/\partial x + y\partial f/\partial y) > z\}$$

Remark 4.2. If sequence (4.34) was constructed on the basis of the matrix (4.16) from Theorem 4.2, then the x-component of the critical point can be evaluated by formula (3.15), where one should set $n := N, c_k := h_{k,0}$, and $s_1 := s_{1,0}$ (i.e. , the first Newton sum of $\mathcal{X}(x)$).

Let us now discuss the significance of the requirements made. The importance of Assumption 4.3 can be realized from the following

EXAMPLE 4.2. *Find the* $\max f$ *for* $f(x,y) := -x^2y^4 - xy^2 - x^2$.

Here the leading form $f_{(6)}(x,y) = -x^2y^4$ does not meet Assumption 4.3, since polynomial $f_{(6)}(1,y) = -y^4$ is only just non-positive but not strictly negative. System (4.30) possesses a single solution, namely $(0,0)$ with $f(0,0) = 0$. However, the *finite* $\sup f(x,y) = 1/4$ is "attained" at *infinity*:

$$f(x,y) - \frac{1}{4} = -x^2 - y^4\left(x + \frac{1}{2y^2}\right)^2 \le 0,\ \lim_{n\to\infty} f\left(\frac{-1}{2n^2}, n\right) = \lim_{n\to\infty}\left(\frac{1}{4} - \frac{1}{4n^4}\right) = \frac{1}{4}.$$

□

As for Assumption 4.6, the situation with its violation is similar to the univariate case considered in Section 3. The condition $\mathcal{D}(\mathcal{F}(z)) = 0$ is equivalent to the existence of two equal critical values for $f(x,y)$. One can no longer claim that the maximal real root of $\mathcal{F}(z)$ coincides with $\max f(x,y)$:

EXAMPLE 4.3. *Find the* $\max f$ *for*

$$f(x,y) := -2x^4 - 16/3x^3y - 4x^2y^2 - 4/3xy^3 - y^4 + 7/6x^2 + xy - 5/6y^2.$$

Here

$$\begin{aligned}\mathcal{F}(z) = 1/21\,743\,271\,936\,\big(&21743271936\,z^9 - 54\,358\,179\,840\,z^8 + 59\,251\,359\,744\,z^7\\ &- 36\,771\,102\,720\,z^6 + 14\,206\,863\,616\,z^5 - 3498\,293\,920\,z^4 + 535\,992\,369\,z^3\\ &- 46\,704\,790\,z^2 + 1771\,561\,z\big) = z(z-11/48)^2(z-1/3)^2(z-11/32)^4.\end{aligned}$$

It turns out that $\max f = f(\pm 1, \mp 1/2) = 11/48$. As for the roots $z = 1/3$ and $z = 11/32$, they correspond to the imaginary critical points:

$$f(\pm i, \mp i) = 1/3;$$

$$f\left(\pm i\frac{2\sqrt{2}+\sqrt{22}}{4}, \mp i\frac{\sqrt{2}+\sqrt{22}}{4}\right) = f\left(\pm i\frac{2\sqrt{2}-\sqrt{22}}{4}, \mp i\frac{\sqrt{2}-\sqrt{22}}{4}\right) = \frac{11}{32}.$$

It should be noted that sequence (4.34)

$$\begin{aligned}H_1(z) &= -1/2(18\,z - 5),\\ H_2(z) &= -5/384\,(288\,z - 107)(18\,z - 5),\\ H_3(z) &= 5/7077\,888\,(288\,z - 107)(7568\,640\,z^2 - 4369176\,z + 611\,435),\\ &\dots\\ H_8(z) &= -6875/36\,028\,797\,018\,963\,968\,(3\,z-1)(48\,z-11)(32\,z-11)^3\\ &\quad\times(71\,912\,448\,z^3 - 41\,525\,760\,z^2 + 5955\,169\,z - 42\,592),\\ H_9(z) &\equiv 3240\,455\,625/1073\,741\,824\,\mathcal{F}(z)\end{aligned}$$

keeps the property discovered for the univariate case: it ignores "extraneous" real roots of $\mathcal{F}(z)$, i.e. those corresponding to the imaginary critical points. For example:

$$P(1, H_1(0.1), \dots, H_9(0.1)) - P(1, H_1(1), \dots, H_9(1)) = 5 - 3 = 2,$$

$$P(1, H_1(0.1), \ldots, H_9(0.1)) - P(1, H_1(0.25), \ldots, H_9(0.25)) = 5 - 3 = 2,$$

and thus, the roots $z = 1/3$ and $z = 11/32$ are not taken into account. On the other hand, this sequence gives the correct information about the critical value $z = 11/48$: it was assumed by the function in *two* real critical points. □

REMARK. The preliminary verification of Assumption 4.6 is unnecessary. It is required only for the case when sequence (4.34) is unable to separate the real roots of $\mathcal{F}(z)$ within a sufficiently large number of substitutions for z.

5. Linear Representation of the Resultant

Theorem 3.3 and part (3) of Theorem 4.3 make a basis for a recurrent method of resultant calculation. Let us sketch the algorithm that gave us the matrices C in Sections 3 and 4. The coefficients c_j of expansion (3.1) for the fraction $G(x)/F(x)$ were the rational symmetric functions of the roots of a polynomial $F(x)$; according to Corollary 3.2, the resultant $\mathcal{R}(F, G)$ can then be found as the determinant of the *Hankel* matrix C (3.5) composed of these coefficients. Since we have obtained such a representation for the univariate resultant, it is possible to find the eliminants $\mathcal{X}(x)$ and $\mathcal{Y}(y)$ for system (2.10) as the appropriate *Hankel* determinants (see formula (4.2)). With these eliminants and the polynomials from the Bézout identities (4.3), (4.4), we can construct expansion (4.10) for the fraction $G(x, y)\mathcal{V}(x, y)/[\mathcal{X}\mathcal{Y}]$. The coefficients of this expansion turn out to be the rational symmetric functions of solutions of system (2.10). Now the resultant $\mathcal{R}(F_1(x, y), F_2(x, y), G(x, y))$ of three polynomials in two variables defined by (2.17) can be computed as the determinant of the block-*Hankel* matrix (4.17) (see formula (4.23)). This allows us to extend the Jacobi method for finding the symmetric functions of solutions of a system (2.18) of three equations in three variables (which works in a similar manner to the bivariate case (von Escherich, 1876; Uteshev and Shulyak, 1992)) and, simultaneously, to make an analogue in $\mathbb{R}^3$ of the Hermite method for the separation of solutions of that system, etc. The leading principal minors of the matrices C play the role of subresultants.

So, we have just climbed the steps of the "shuttle" procedure for the recurrent resultant computation that was outlined in Section 2. One important step must be fixed, however. To extend the Jacobi method to $\mathbb{R}^3$, a linear representation for the resultant $\mathcal{R}(F_1, F_2, G)$ is needed. This will be given by the following

THEOREM 5.1. *Let $m := \deg G \le n := \deg F_1 = \deg F_2$. Suppose that conditions of the Theorem* 4.3 *are satisfied, and $C_N \neq 0$. There exist polynomials $\mathcal{M}_1(x, y), \mathcal{M}_2(x, y)$ and $\mathcal{Q}_1(x, y)$ providing the fulfilment of the identity*

$$\mathcal{M}_1(x, y)F_1(x, y) + \mathcal{M}_2(x, y)F_2(x, y) + \mathcal{Q}_1(x, y)G(x, y) \equiv C_N/d_{2n-2,0} \tag{5.1}$$

and of the degrees satisfying the following restrictions

$$\begin{cases} \deg \mathcal{M}_j = \deg_x \mathcal{M}_j \le n + m - 2, & \deg_y \mathcal{M}_j \le n - 1; (j = 1, 2) \\ \deg \mathcal{Q}_1 = \deg_x \mathcal{Q}_1 \le 2n - 2, & \deg_y \mathcal{Q}_1 \le n - 1 \end{cases} \tag{5.2}$$

PROOF. Let us rearrange the rows and columns of matrix C, ordering the monomials of the basis (4.20) according to the magnitude of their degrees. For simplicity, denote again

the new matrix by C. Find the expressions for F_1 and F_2 from equalities (4.3) and (4.4) :

$$F_1 \equiv (\mathcal{Q}\mathcal{X} - \mathcal{N}\mathcal{Y})/\mathcal{V}, \quad F_2 \equiv (\mathcal{M}\mathcal{Y} - \mathcal{P}\mathcal{X})/\mathcal{V}.$$

Substitute these in (5.1):

$$M_x(x,y)\mathcal{X}(x) + M_y(x,y)\mathcal{Y}(y) + \mathcal{Q}_1(x,y)G(x,y)\mathcal{V}(x,y) \equiv \mathcal{V}(x,y)C_N/d_{2n-2,0}. \quad (5.3)$$

Here

$$M_x := \mathcal{M}_1\mathcal{Q} - \mathcal{M}_2\mathcal{P}, \quad M_y := \mathcal{M}_2\mathcal{M} - \mathcal{M}_1\mathcal{N}. \quad (5.4)$$

According to the restrictions (5.2) and (4.5), M_x and M_y have to satisfy the conditions

$$\begin{cases} \deg M_x \le N+m-2,\ \deg_x M_x \le 2n+m-3,\ \deg_y M_x \le N-1; \\ \deg M_y = \deg_x M_y \le N+m-2,\ \deg_y M_y \le 2n-2. \end{cases} \quad (5.5)$$

Divide the equality (5.3) by the product $[\mathcal{X}(x)\mathcal{Y}(y)]$ and expand both its sides in the series in the negative powers of x and y (using the expansions (4.7), (4.10), and taking into account the condition (4.8)):

$$\frac{M_x(x,y)}{\mathcal{Y}(y)} + \frac{M_y(x,y)}{\mathcal{X}(x)} + \mathcal{Q}_1(x,y)\left[L_1(x,y) + \sum_{j,k=0}^{\infty} \frac{c_{jk}}{x^{j+1}y^{k+1}}\right] \equiv \frac{C_N}{d_{2n-2,0}} \sum_{\substack{j,k=0 \\ j+k\ge 2n-2}}^{\infty} \frac{d_{jk}}{x^{j+1}y^{k+1}}. \quad (5.6)$$

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 the 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 (5.6). On the right-hand side they all vanish except for that of $x^{-5}y^{-1}$, the latter equals C_9. So, one gets a linear system for the coefficients of $\mathcal{Q}_1(x,y)$:

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

Solving this by Cramer's rule, we obtain the following result: the polynomial $\mathcal{Q}_1(x,y)$ from the equality (5.1) equals the determinant of the matrix obtained on replacing the last row in the matrix C by the row of monomials of the basis (4.20) (where one should replace $\alpha \to x,\ \beta \to y$).

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 univariate case (see proof of Theorem 3.3).

To find now $M_y(x,y)$, let us consider it as a polynomial in y and exploit the equality (5.3). Equate the coefficients of $y^{N+2n-2}, \ldots, y^N$ on both its sides. Because of the restrictions on the degrees (5.5) and (4.6), only these coefficients which are contained in the summands $M_y\mathcal{Y}$ and $\mathcal{Q}_1G\mathcal{V}$ differ from zero. 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 (5.3). Finally, $\mathcal{M}_1(x,y)$ and $\mathcal{M}_2(x,y)$ can be found from (5.4).

Another method for finding $\mathcal{M}_1$ and $\mathcal{M}_2$ can be proposed after obtaining the expression for $\mathcal{Q}_1(x,y)$. Consider the equality (5.1) rewritten as

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

together with (4.4): $\mathcal{M}F_1 + \mathcal{N}F_2 \equiv \mathcal{X}(x)$. Then one obtains

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

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

Polynomials and the division procedure in brackets are considered with respect to the variable y. One may notice an analogy with the remark at the end of the proof of Theorem 3.3. □

The idea of constructing the resultant as the block Hankel matrix can be found in Laurent (1900, p.42). This interesting book contains many good ideas between numerous errors, with a few (but not all) of them mentioned in Macaulay (1903, p.4). There exists a relationship between the polynomials of the linear representation of the resultant (5.1) (if we construct them for the basis (4.25) via the procedure used in the proof of Theorem 5.1) and similar ones from Sections 6–8 of the book by Macaulay (1916): the multipliers of $G(x,y)$ are the same (up to a constant), but those of $F_1(x,y)$ and $F_2(x,y)$ are different.

6. The Constrained Maximum

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

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

where the constraint set $\mathbb{S}$ is defined as that of the real solutions of the system of algebraic inequalities

$$G_1(X) \geq 0, \ldots, G_p(X) \geq 0. \tag{6.1}$$

In computer algebra this problem can be solved, for example, by using a cylindrical algebraic decomposition algorithm applied to the set $\{X \in \mathbb{R}^K \mid f(X) - z \geq 0, G_1(X) \geq 0, \ldots, G_p(X) \geq 0\}$ (Bank *et al.* (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 formula

$$\begin{aligned}
&\operatorname{nrs}\{(2.10) \mid G_1 > 0, \ldots, G_p > 0\} \\
&= \frac{1}{2^{p-1}}[(1 - 2^{p-1}) \operatorname{nrs}\{(2.10)\} + \sum_{1 \leq j_1 \leq p} \operatorname{nrs}\{(2.10) \mid G_{j_1} > 0\} \\
&+ \sum_{1 \leq j_1 < j_2 \leq p} \operatorname{nrs}\{(2.10) \mid G_{j_1} G_{j_2} > 0\} + \cdots + \operatorname{nrs}\{(2.10) \mid G_1 G_2 \ldots G_p > 0\}]
\end{aligned} \tag{6.2}$$

established first by Markov (1940). Although the well-known results by Ben-Or *et al.*

(1986) permit one to compute this number in a more efficient way, this formula will be enough to explain the idea of our approach.

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

ASSUMPTION 6.1. *We shall assume the domain $G(x, y) \geq 0$ to be bounded.*

This can be achieved by imposing on polynomial G the restrictions similar to those made in Section 4 for the function f: either Assumptions 4.3 or 4.4. In particular, let $m := \deg G$ be even. As for the polynomial f, we now do not impose any restriction on it.

The 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

$$G(x, y) = 0, \quad \mathcal{J}(G, f) := \partial G/\partial x \cdot \partial f/\partial y - \partial f/\partial x \cdot \partial G/\partial y = 0. \tag{6.3}$$

So, we have to compare the two values

$$z_> := \max_{G(x,y)>0} f(x, y), \quad \text{and} \quad z_= := \max_{G(x,y)=0} f(x, y) \tag{6.4}$$

As in Section 4 consider then

PROBLEM 6.1. For every real z find

$$\mathrm{nrs}\{(4.30) \mid G(x, y) > 0, f(x, y) > z\} \quad \text{and} \quad \mathrm{nrs}\{(6.3) \mid f(x, y) > z\} \tag{6.5}$$

Under the assumptions of Theorem 4.2 or 4.3, one can find the second of these numbers with the help of an appropriate block Hankel matrix. The sequence of its leading principal minors

$$H_1^*(z), \ldots, H_{m(m+n-1)}^*(z) \tag{6.6}$$

permits one to separate the roots of $H_{m(m+n-1)}^*(z)$ which coincide with the values of f at the solutions of (6.3). Thus, one may localize $z_=$, i.e. the maximum of f at the boundary of the set.

Now to find the $\mathrm{nrs}\{(4.30) \mid G > 0, f > z\}$ one has to construct, along with the sequence (4.34), the third Sturm series

$$H_1^{**}(z), \ldots, H_N^{**}(z) \tag{6.7}$$

for evaluating $\mathrm{nrs}\{(4.30) \mid (f - z)G > 0\}$. By Corollaries 4.1 and 4.2, the polynomial $H_N^{**}(z)$ differs from $H_N(z)$ (and thus, from $\mathcal{F}(z)$ introduced by (4.36)) only by the constant factor. The use of the formula (6.2) gives us the following equality:

$$\begin{aligned} \mathrm{nrs} \quad & \{(4.30) \mid G > 0, a < f < b\} \\ & = \tfrac{1}{2}[P(1, H_1(a), \ldots, H_N(a)) - P(1, H_1(b), \ldots, H_N(b) \\ & \quad + P(1, H_1^{**}(a), \ldots, H_N^{**}(a)) - P(1, H_1^{**}(b), \ldots, H_N^{**}(b)] \end{aligned} \tag{6.8}$$

which provide us with an opportunity to localize the root of $\mathcal{F}(z)$ coinciding with $z_>$, i.e. the maximum of the objective function inside the constraint set, and to compare it with $z_=$.

EXAMPLE 6.1. *Find $\max_{G(x,y)\geq 0} f(x, y)$ for the polynomial f from Example* 4.1 *and for* $G := -x^2 - y^2 - xy + 10x + 8y - 27$.

The sequence (6.6):

$$
\begin{aligned}
H_1^*(z) &= -163\,783\,454\,148\,z + 53\,177\,999\,662\,367,\\
H_2^*(z) &= 33\,043\,355\,481\,703\,666\,905\,z^2 - 21\,382\,843\,070\,791\,551\,797\,361\,z\\
&\quad +3459\,158\,803\,313\,240\,668\,027\,318,\\
&\ldots\\
H_8^*(z) &= 9949\,844\,839\,491\,z^8 - 25\,844\,507\,835\,910\,362\,z^7 + 29\,350\,431\,452\,148\,112\,773\,z^6\\
&\quad -19\,034\,022\,603\,127\,214\,991\,852\,z^5\\
&\quad +7709\,490\,826\,442\,511\,053\,846\,733\,z^4\\
&\quad -1997\,055\,838\,529\,139\,670\,618\,667\,978\,z^3\\
&\quad +323\,082\,798\,240\,279\,672\,149\,109\,601\,675\,z^2\\
&\quad -29\,844\,710\,775\,253\,389\,113\,553\,491\,156\,624\,z\\
&\quad +1205\,188\,271\,972\,045\,186\,872\,415\,235\,073\,280
\end{aligned}
$$

permits one to isolate the second of the numbers (6.4): $z_{=} = 349.61210 \pm 10^{-5}$.

Let us now evaluate the number $z_{>}$. To compute the entries of the matrix H^{**} for $\operatorname{nrs}\{(4.30) \mid (f-z)G > 0\}$ one may use the computations made for Example 4.1. Indeed, the formula is similar to (4.11):

$$h_{jk}^{**} = \sum_{\substack{p,\ell=0\\ p+\ell\le m}} b_{p\ell} h_{p+j,\ell+k}$$

with h_{jk} taken from that example and $G(x,y)$ defined by (4.1).

$$
\begin{aligned}
H_1^{**}(z) &= 902\,375\,z - 1591\,750\,999,\\
H_2^{**}(z) &= 243\,546\,139\,125\,z^2 - 601\,348\,448\,837\,898\,z + 221\,634\,739\,662\,079\,765,\\
&\ldots\\
H_9^{**}(z) &= \frac{68\,696\,410\,531\,813\,777\,559\,965\,184\,896\,794\,624}{1220\,703\,125} H_9(z)
\end{aligned}
$$

with $H_9(z)$ computed in Example 4.1. By formula (6.8), there exist exactly three critical points of $f(x,y)$ lying inside the domain. Two of the corresponding critical values lie inside $]356.1, 356.3[$, and the remaining one gives the value of the maximum

$$z_{>} := \max_{G(x,y)>0} f(x,y) = \max_{G(x,y)\ge 0} f(x,y) = 361.36917 \pm 10^{-5}.$$

The corresponding critical point can be restored via formulae (4.38):

$$(4.46410 \pm 10^{-5}, 1.73205 \pm 10^{-5})$$

(with the exact position at $(1 + 2\sqrt{3}, \sqrt{3})$). □

Remark. The numbers (6.5) are not independent. According to the Kronecker–Poincaré index theorem (Uteshev, 1991), there exists a relationship between the number of critical points of $f(x,y)$ inside the domain $G(x,y) \ge 0$ and on its boundary. If, for example, the curve $G(x,y) = 0$ consists of a single oval, then:

$$
\begin{aligned}
&\operatorname{nrs}\{\,(4.30) \mid G > 0, \mathcal{H}(f) > 0\,\} - \operatorname{nrs}\{\,(4.30) \mid G > 0, \mathcal{H}(f) < 0\,\}\\
&= 1 + 1/2\big(\operatorname{nrs}\{\,(6.3) \mid \mathcal{J}(\mathcal{J}(G,f),f) > 0\,\} - \operatorname{nrs}\{\,(6.3) \mid \mathcal{J}(\mathcal{J}(G,f),f) < 0\,\}\big)\,.
\end{aligned}
$$

Here $\mathcal{H}(f)$ is the Hessian of f (4.35), and the first term on the left-hand side gives the number of extremal points (local maxima and minima) of f inside the curve.

Another approach for finding the first of the numbers (6.5) consists of transforming the inequality $G(x,y) > 0$ into a new one with respect to the variable z. The latter is possible since, under Assumption 4.6, the roots of $\mathcal{F}(z)$ are rationally connected with the coordinates of corresponding critical points :

$$x = \theta(z), \quad y = \xi(z)$$

(analogues of formulae (4.38) from Example 4.1). Denote $\mathcal{G}(z) := G(\theta(z), \xi(z))$, then

$$\mathrm{nrs}\{(4.30) \mid G(x,y) > 0, a < f(x,y) < b\} = \mathrm{nrr}\{\mathcal{F}(z) = 0 \mid \mathcal{G}(z) > 0, a < z < b\}.$$

In this way, the multivariate optimization problem can be reduced to the univariate separation one.

The approach proposed can be extended to the general case when the constraint set $\mathbb{S}$ is given by the system of inequalities (6.1). In addition to the systems (4.30) and (6.3), another type of systems appears: $G_q(x,y) = 0, G_\ell(x,y) = 0$ yielding the vertices ("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 (6.1) can be treated itself as the constrained optimization problem (Uteshev and Cherkasov, 1996). So, for example, under Assumption 6.1, the set defined by $G(x,y) \geq 0$ is not empty iff

$$\mathrm{nrs}\{\partial G/\partial x = 0, \partial G/\partial y = 0 \mid G > 0\} > 0.$$

Moreover, on this way one can even establish the geometry of its boundary:

THEOREM 6.1. (PETROWSKY, 1938) *The difference between the number of positive and the number of negative ovals of the curve $G = 0$ equals*

$$\mathrm{nrs}\left\{\frac{\partial G}{\partial x} = 0, \frac{\partial G}{\partial y} = 0 \;\middle|\; G > 0, \mathcal{H}(G) > 0\right\} - \mathrm{nrs}\left\{\frac{\partial G}{\partial x} = 0, \frac{\partial G}{\partial y} = 0 \;\middle|\; G > 0, \mathcal{H}(G) < 0\right\}. \tag{6.9}$$

Here $\mathcal{H}(G)$ is the Hessian of G.

EXAMPLE 6.2. *Find an estimation for the number of ovals of the curve $G(x,y) := f(x,y) = 0$, where f is the polynomial from Example* 4.1.

With the help of the Markov's formula (6.2), the difference (6.9) can be transformed into

$$(\sigma(H^\star) + \sigma(H^{\star\star}))/2$$

with $H^\star$ (or $H^{\star\star}$) being the matrix (4.18) built for the system (4.30) and for $G(x,y) := \mathcal{H}(f)$ (or for $G(x,y) := \mathcal{H}(f) \cdot f$). Calculating the numbers of permanences and variations in the corresponding sequences of the leading principal minors, we evaluate the difference (6.9): it equals $(1+5)/2 = 3$. Because of the Harnack inequality (Petrowsky, 1938), the total number of ovals of the curve of the order m does not exceed $(m-1)(m-2)/2+1$. For our example, this leads to an unambiguous deduction: the curve of interest has exactly three positive and no negative ovals. □

7. Conclusions

We have treated here the problem of finding the maximum value of a polynomial in an algebraic domain. Using Hermite's method of the separation of the solutions of an algebraic system, we reduce the stated problem to a univariate one, investigating the critical values of the polynomial. The proposed approach is free from the usual assumptions on the convexity of the objective function or the constraint set.

We have also discussed the possibility of unifying the separation and elimination algorithm for a polynomial equation system, using the triple Hankel or block Hankel matrices. Matrices of these very special structures are known to have several nice properties from the computational point of view (Iohvidov, 1982; Bini and Pan, 1994).

Our investigation can, by no means, be considered as complete. We note just a few problems for further investigation:

Detailed comparison of the proposed *Hankel* approach for the resultant computation with the other algorithms (like the determinantal one (González-Vega, 1991) or those based on the Gröbner basis construction) has to be discussed;

It would be interesting to establish the structure of Assumption 4.6 in terms of the function $f(x, y)$ (i.e. to find an analogue of Theorem 3.5 for the bivariate case);

The computational complexity of the algorithm has to be estimated:

Although the constraint optimization problem was solved in Section 6 by a purely algebraic method, it seems desirable to exploit it together with some hybrid ones, e.g. those using interval arithmetic (Collins and Krandick, 1993). Its employment might be advantageous for optimization problems with objective functions or constraint sets depending on parameters.

We hope to discuss these problems in subsequent papers.

Acknowledgements

The research described in this publication was made possible in part by Grant No. JKF100 from the International Science Foundation and Russian Government, and grant from ÖAD GZ 560.302/1-IV/A/5a/94. The final version was prepared when the second author was a Soros Graduate Student Grantee, grant A97-1157.

We would like to thank the referees for many constructive suggestions which helped to improve the presentation.

References

Akritas, A.G. (1989). *Elements of Computer Algebra with Applications*. New York: Wiley.

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

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. Comput. Sys. Sci.* **32**, 251–264.

Bini, D., Pan, V. (1994). *Polynomial and Matrix Computations*, volume 1. Boston: Birkhäuser.

Collins, G.E. (1971). The calculation of multivariate polynomial resultants. *J. ACM* **18**, 515–532.

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

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

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

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

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

Iohvidov, I.S. (1982). *Hankel and Toeplitz Matrices and Forms*, Boston: Birkhäuser.

Jury, E.I. (1974). *Inners and Stability of Dymamic Systems*, New York: Wiley.

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**, 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.

Macaulay, F.S. (1916). *The Algebraic Theory of Modular Systems*, Cambridge: Cambridge University Press.

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

Netto, E. (1896–1900), *Vorlesungen über Algebra*, vol. 2. Teubner, Leipzig.

Pedersen, P. (1991). Calculating multidimensional symmetric functions using Jacobi's formula. In *Proc. 9th Int. Symp., AAECC-9, Lecture Notes Computer Science* **539**, pp. 304–317. Berlin: Springer.

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

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

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

Uteshev, A. Yu. (1991). Calculation of the Kronecker–Poincaré index of an algebraic field with respect to an algebraic curve in the plane (in Russian). *Differentsial'nye Uravnenija* **27**, 2181–2183.

Uteshev, A.Yu., Cherkasov, T.M. (1996). The localization of solutions to systems of algebraic equations and inequalities: the Hermite method. *Doklady Mathematics* **53**, 227–229.

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.

von Escherich, G. (1876). Beitrage zur Bildung der symmetrischen Functionen der Wurzelsysteme und der Resultante simultaner Gleichungen. *Denkschriften, Abt.2, Österr. Akad. Wiss. Math.-Naturwiss. Klasse* **36**, 251–272.

Originally received 7 September 1994
Accepted 2 August 1997