J. Symbolic Computation (1990) 10, 1–5

Testing Polynomials

JEAN-JAQUES RISLER† AND FELICE RONGA‡

† *Université P. et M. Curie, Laboratoire d'analyse numérique, 4, place Jussieu, F-75230 Paris Cedex 05, France*

‡ *Section de Mathématiques, Université de Genève, 2–4 rue du Lièvre, case postale 240, CH-1211 Genève 24, Switzerland*

(*Received 29 June 1987*)

Let $S \subset \mathbb{N}^n$ be a finite set $(\alpha_1, \ldots, \alpha_k)$ of exponents. We construct explicitly a testing set $T_S \subset \mathbb{N}^n$ with k elements $t_1, \ldots, t_k$ (namely $t_i = (2^{\alpha_i^1}, \ldots, 2^{\alpha_i^n})$), such that if

$$P = \sum_{\alpha \in S} a_\alpha X^\alpha \in \mathbb{Q}[X_1, \ldots, X_n],$$

then there exists i $(1 \le i \le k)$ such that $P(t_i) \neq 0$.

1. Introduction

We thank the referee for suggesting to us the following introductory comment:

"Many algorithms in computer algebra which deal with sparse polynomials make the assumption that a polynomial which evaluates to zero is identically zero. The purpose of this paper is to explore the question: How do we guarantee that a polynomial is zero? We show that any set of evaluations (a *testing set*) for a polynomial with k terms has to contain at least k members, and we construct such a set with precisely k members. This construction relies on knowing the exponents, but not the coefficients, of the polynomial being tested. We leave to others the task of converting this theorem into algorithms."

More precisely, let S be a finite subset of $\mathbb{N}^n$, i.e. S is a finite collection $\{\alpha_1, \ldots, \alpha_k\}$ of distinct multi-indices $\alpha_i = (\alpha_i^1, \ldots, \alpha_i^n)$, where α_i^h is a non-negative integer. Let $\mathbb{K}$ be a field of zero characteristic (e.g. $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$) and denote by P_S the vector space of polynomials of type S, that is the set of polynomials of the form

$$\sum_{\alpha \in S} a_\alpha X^\alpha$$

with $a_\alpha \in \mathbb{K}$, where $X = (X_1, \ldots, X_n)$ and $X^\alpha = X_1^{\alpha^1} \ldots X_n^{\alpha^n}$.

Buscar ejemplos para caracterizar que no son constantes.

DEFINITION. A set of points T_S of $\mathbb{K}^n$ is said to be a testing set for S if for any $P \in P_S$, $P|T_S = 0$ implies that $P \equiv 0$.

To our knowledge, the question of finding testing sets has been raised by J. H. Davenport and B. Trager. We will prove the following:

THEOREM 1.1. *Let $S = \{\alpha_1, \ldots, \alpha_k\}$ be a finite subset of N^n and set $T_i = (2^{\alpha_i^1}, \ldots, 2^{\alpha_i^n}) \in \mathbb{K}^n$ and $T_S = \{T_1, \ldots, T_k\} \subset \mathbb{K}^n$. Then T_S is a testing set for S. Any testing set must contain at least k elements.*

0747-7171/90/070001+05 $03.00/0

In other words, if P is not identically zero, then $P(T_i) \neq 0$ for some i, and T_S is optimal in the sense that it could not contain less points. As we shall see, in the definition of T_S the number 2 can be replaced by any rational number (or real number if $\mathbb{K} \subset \mathbb{R}$) greater than 1.

DEFINITION. Let δ be a direction in $\mathbb{K}^n$ and let H be an affine hyperplane in $\mathbb{K}^n$ not parallel to δ; denote by $\pi_\delta : \mathbb{K}^n \to H$ the projection parallel to δ. We shall say that δ is generic with respect to the hypersurface $V(P) = \{x \in \mathbb{K}^n | P(x) = 0\}$, where P is some polynomial, if $\pi_\delta | V(P)$ is a finite map. This amounts to say that for any line l parallel to δ, $l \cap V(P)$ is a finite set.

The following is an easy consequence of the theorem:

COROLLARY. *Let $P = \sum_{\alpha \in S} a_\alpha X^\alpha$ be a polynomial and set $d_\infty = \sup\{|\alpha|, a_\alpha \neq 0\}$, where $|\alpha| = \alpha_1 + \ldots + \alpha_n$, and $S_\infty = \{\alpha \in S | |\alpha| = d_\infty\}$. Then among the directions of the vectors of T_{S_∞} there is one that is generic for P.*

Note that every element of a T_S is non-zero and therefore defines a direction in $\mathbb{K}^n$.
We are grateful to Pierre de la Harpe for helpful conversations on the Schur product.

2. Proof of the Theorem

Let $P = \sum_{\alpha \in S} a_\alpha X^\alpha$ be a polynomial of type S and let $Y_S = \{Y_1, \ldots, Y_k\}$, where $Y_i = (Y_i^1, \ldots, Y_i^n)$, be a set of k points in $\mathbb{K}^n$. Consider the system of equations:

$$\sum_{\alpha \in S} a_\alpha Y_i^\alpha = 0, \quad i = 1, \ldots, k.$$

We can view them as a system of k linear equations with k unknowns a_α, $\alpha \in S$, and therefore our theorem will be a consequence of the following proposition:

PROPOSITION 2.1. *Let $a \in R$, $a > 1$ and $S = \{\alpha_1, \ldots, \alpha_k\} \subset N^n$ satisfy $\alpha_i \neq \alpha_j$ for $i \neq j$ and set $T_i = (a^{\alpha_i^1}, \ldots, a^{\alpha_i^n})$. Then the determinant of the matrix $(T_i^{\alpha_j})_{i=1\ldots k, j=1\ldots k}$ is different from zero.*

The proof will be split into several lemmas. For X and Y in $\mathbb{R}^n$ we shall write

$$\langle X, Y \rangle = \sum_{i=1\ldots n} X^i \cdot Y^i$$

and so $T_i^{\alpha_j} = a^{\langle \alpha_i, \alpha_j \rangle}$. The matrix

$$M = M(a, \alpha_1, \ldots, \alpha_k) = (a^{\langle \alpha_i, \alpha_j \rangle})_{i=1\ldots k, j=1\ldots k}$$

is symmetric.

LEMMA 2.2. *Let $S = \{\alpha_1, \ldots, \alpha_k\} \subset \mathbb{R}^n$ be a set of k distinct points. Then there is an orthogonal transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ arbitrarily close to the identity map such that if we set $\beta_i = A(\alpha_i)$, then $\beta_i^h \neq \beta_j^h$ for any $h = 1, \ldots, n$ and $i, j = 1, \ldots, k$, $i \neq j$.*

PROOF. If $l \subset \mathbb{R}^n$ is an affine line and H a hyperplane in $\mathbb{R}^n$, then $A(l)$ will not be contained in H for almost all orthogonal $A : \mathbb{R}^n \to \mathbb{R}^n$. Let $L = \{$lines joining pairs of distinct vector of

$S\}$, a finite set of lines, and set $H_i = \{X_i = 0\}$, $i := 1, \ldots, n$. Then for almost all A we will have $A(l) \notin H_i$, for any $l \in L$ and $i = 1, \ldots, n$. Such an A has exactly the properties we are looking for.

DEFINITION. Let $A = (a_{i,j})$ and $B = (b_{i,j})$ be two matrices of the same size with coefficients in any ring. The Schur product $A * B$ of A and B is the matrix defined by:

$$(A * B)_{i,j} = a_{i,j} \cdot b_{i,j}.$$

Note that we will write $A \cdot B$ for the ordinary matrix product. Recall that a symmetric matrix with real coefficients is said to be positive if $\langle A(v), v\rangle \geqslant 0$ for all v in $\mathbb{R}^n$. It is said to be positive-definite if in addition $\langle A(v), v\rangle = 0$ implies $v = 0$. Clearly, if A is positive, it is positive-definite if and only if $det(A) \neq 0$. Also, if A and B are positive then so is $A + B$, and if one of the two is definite then so also is $A + B$.

LEMMA 2.3 (cf. Schur, 1911, section VII). *Let A and B be two $k \times k$ symmetric matrices.*

(i) If A and B are positive then so is $A * B$.
(ii) If A and B are positive-definite then so is $A * B$.

PROOF. Since B is positive, there exists a symmetric $k \times k$ matrix $C = (c_{i,j})$ such that $B = C \cdot C$ (ordinary matrix product here!). Therefore

$$b_{i,j} = \sum_h c_{i,h} \cdot c_{h,j}$$

and so

$$\langle A * B(v), v\rangle = \sum_{i,j} a_{i,j} \cdot b_{i,j} \cdot v^i \cdot v^j = \sum_{h,i,j} a_{i,j} \cdot c_{i,h} \cdot c_{j,h} \cdot v^i \cdot v^j$$
$$= \sum_{h,i,j} a_{i,j} \cdot w_h^j \cdot w_h^i = \sum_h \langle A(w_h), w_h\rangle,$$

where $w_h = (c_{i,h} \cdot v^i)_{i=1\ldots k}$. Since A is positive, the last expression is non-negative. If A and B are definite, then so is C because $det(B) = det(C)^2$; therefore $A * B(v), v\rangle = 0$ implies $A(w_h) = 0$, for all $h = 1 \ldots k$, which implies $w_h = 0$ and so $c_{i,h} \cdot v^i = 0$ for all h, i. But for each i there must be an h such that $c_{i,h} \neq 0$, otherwise $det(C) = 0$, and so $v^i = 0$ for all $i = 1 \ldots k$.

DEFINITION. The Schur exponential of the $k \times k$ symmetric matrix A is defined by:

$$e^{*A} = \sum_{n=0\ldots\infty} (1/n!) \cdot A^{*n}$$

where $A^{*n} = A * \cdots * A$, n times.

In other words, $(e^A)_{i,j} = e^{a_{i,j}}$. Since e^{*A} is a linear combination of Schur products of A with positive coefficients, it follows from Lemma 2.3 that it is positive (or positive-definite) if A is. Note that e^{*0} is the matrix with all entries equal to 1.

LEMMA 2.4. *Let $b_1, \ldots, b_k$ be distinct real numbers and set $B = (\mu \cdot b_i \cdot b_j)_{i,j=1\ldots k}$, where μ is a strictly positive real number. Then e^{*B} is positive-definite.*

PROOF. Let $b = \sqrt{\mu} \cdot (b_1, \ldots, b_k)$; then $B = {}^t b \cdot b$ (ordinary matrix product) and so

$$\langle B(v), v\rangle = \langle {}^t b \cdot b(v), v\rangle = \langle b(v), b(v)\rangle \geq 0.$$

Therefore, by Lemma 2.3, B^{*h} is positive for all $h \geq 0$ and it is enough to prove that $\sum_{h=0\ldots k-1} B^{*h}$ is positive-definite. If, indeed, $v \in \mathbb{R}^n$ is such that $B^{*h}(v) = 0$ for $h = 0 \ldots k-1$, then $\langle b^{*h}(v), b^{*h}(v)\rangle = 0$ and hence $b^{*h}(v) = 0$, $h = 0 \ldots k-1$. Written in matrix notation:

$$\begin{pmatrix} 1 & \ldots & 1 \\ b_1 & \ldots & b_k \\ \vdots & \ddots & \vdots \\ b_1^{k-1} & \ldots & b_k^{k-1} \end{pmatrix} \cdot v = 0$$

and the determinant of the matrix on the left is the well-known Van der Monde determinant, and equals $\prod_{1 \leq i < j \leq n} (b_i - b_j)$; it is therefore different from zero, and hence $v = 0$.

PROOF OF THE PROPOSITION. By Lemma 2.2 there is an orthogonal transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ such that all the co-ordinates of the vectors $A(\alpha_1), \ldots, A(\alpha_k)$ are different. Now

$$a^{\langle \alpha_i, \alpha_j\rangle} = e^{Ln(a) \cdot \langle \alpha_i, \alpha_j\rangle}$$

and therefore

$$M(a, \alpha_1, \ldots, \alpha_k) = e^{*(Ln(a) \cdot \langle \alpha_i, \alpha_j\rangle)} = e^{*(Ln(a) \cdot \langle A(\alpha_i), A(\alpha_j)\rangle)}.$$

Set $\beta_i = A(\alpha_i)$. We have

$$e^{Ln(a) \cdot (\sum_h \beta_i^h \cdot \beta_j^h)} = \prod_{h=1,\ldots,n} e^{Ln(a) \cdot \beta_i^h \cdot \beta_j^h}$$

and so

$$M(a, \alpha_1, \ldots, \alpha_k) = e^{M^1} * \cdots * e^{M^n},$$

where

$$M^h = (Ln(a) \cdot \beta_i^h \cdot \beta_j^h)_{i,j=1\ldots k}.$$

By Lemma 2.4, $e^{*Ln(a) \cdot M^h}$ is positive-definite for $h = 1, \ldots, n$ and therefore $M(a, \alpha_1, \ldots, \alpha_k)$ also, by Lemma 2.3.

PROOF OF THE COROLLARY. Let $P_\infty = \sum_{\alpha \in S_\infty} a_\alpha X^\alpha$ be the homogeneous part of highest degree of P. If $l = \{t \cdot v + w | t \in \mathbb{K}\}$ is a line on which P vanishes, then the polynomial $\phi(t) = t^{d_\infty} \cdot P(v/t + w)$ is identically zero, and $\phi(0) = P_\infty(v)$. The corollary follows now from the theorem.

Nota: Un polynomio se hace zero en un numero finito de puntos o en una linea completa!

3. Some Remarks and Comments

(a) In the case of polynomials in one variable, say with real or complex coefficients, Theorem 1.1 is a consequence of the Descartes' well-known lemma, which implies that for any sequence $(\alpha_i)_{i=1\ldots k} \subset \mathbb{R}_+$, $\alpha_1 \leq \cdots \leq \alpha_k$, if $P(\alpha_i) = 0$ for $i = 1 \ldots k$, then $P \equiv 0$:

DESCARTES' LEMMA. *Let $P \in \mathbb{R}[X]$ be a polynomial in one variable with k non-zero coefficients; then P has at most $k-1$ strictly positive real roots.*

The proof is easy: use induction on k and Rolle's theorem.

(b) In our first attempt to construct a testing set for a given $S = \{\alpha_1, \ldots, \alpha_k\}$ we took simply $l_i = \alpha_i$. In view of Proposition 2.1, this is a testing set if and only if $det(\alpha_i^{\alpha_j}) \neq 0$, where for α and β in $\mathbb{N}^n$ we set $\alpha^\beta = (\alpha^1)^{\beta^1} \cdot \ldots \cdot (\alpha_n)^{\beta^n}$ and $0^0 = 1$.

CONJECTURE. *If* $\alpha_1, \ldots, \alpha_k \in \mathbb{N}^n$ *are distinct sequences of non-negative integers, then* $det(\alpha_i^{a_j}) \neq 0$.

REMARKS.

(1) It follows from Descartes' lemma that for $n = 1$ the conjecture is true.

(2) The conjecture is true for $k = 2$: $det(\alpha_i^{a_j}) = 0$ implies

$$(\alpha_1^1)^{a_1} \cdot \ldots \cdot (\alpha_1^n)^{a_1^n} \cdot (\alpha_2^1)^{a_2^1} \cdot \ldots \cdot (\alpha_2^n)^{a_2^n} - (\alpha_1^1)^{a_2^1} \cdot \ldots \cdot (\alpha_1^n)^{a_2^n} \cdot (\alpha_2^1)^{a_1^1} \cdot \ldots \cdot (\alpha_2^n)^{a_1^n} = 0.$$

But if a and b are non-negative reals, then $a^a \cdot b^b - a^b \cdot b^a \geq 0$ and equality holds exactly when $a = b$. From this and the above equation it follows easily that $\alpha_1 = \alpha_2$.

(c) Theorem 1.1 asserts that for a set S of k multi-indices any testing set must contain at least k points. One may ask for analogous bounds for the number of tests in other settings. For instance, let C_k be the set of polynomials $P \in \mathbb{R}[X_1, \ldots, X_n]$ of additive complexity $C_+^{\mathbb{R}} \leq k$ (see Risler (1985) for details on the additive complexity).

It is proved in Risler (1985) that if $P \in \mathbb{R}[X]$ and $P \in C_k$, then the number of real roots of P does not exceed 5^{k^2}. This implies that for a polynomial $P \in C_k$, if $\alpha_i \in \mathbb{R}$, $i = 1 \ldots N$, are such that $\alpha_1 < \cdots < \alpha_N$ and $N = 5^{k^2} + 1$, then there exists i such that $P(\alpha_i) \neq 0$. This can be easily generalised to polynomials in n variables.

PROBLEM. Find an explicit testing set for C_k.

References

Risler, J.-J. (1985). Additive complexity of real polynomials. *S.I.A.M. Journal of Computing* **14**, 1.

Schur, J. (1911). Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veräanderlichen. *Journal für die reine und angewandte Mathematik* **140**, 1–28.

J. Symbolic Computation (1990) **10**, 39–51

Complexity of Computation on Real Algebraic Numbers

MARIE-FRANÇOISE ROY† AND AVIVA SZPIRGLAS‡

† *Institut de Recherche Mathématique de Rennes, Université de Rennes I, Campus de Beaulieu, F-35042 Rennes Cedex, France*

‡ *CSP, Université Paris Nord, F-93430 Villetaneuse, France*

(*Received* 14 *January* 1988)

This paper is devoted to a precise algorithmical and complexity study of a new polynomial time method for formal computations with polynomial inequalities and real algebraic numbers.

1. Introduction

A new method for coding the real algebraic numbers, based on the use of Thom's lemma and the study of simultaneous inequalities from Ben-Or, Kozen & Reif (1986), has been introduced by Coste & Roy (1988). This leads to various applications in the field of computational real algebraic geometry: study of the topology of a real algebraic curve (Roy, 1987), or of the analytic branches of a real algebraic curve (Cucker *et al.*, 1987).

In this paper, we give improved versions of the algorithms in Coste & Roy (1988) and we study their complexity.

In the second section we introduce some basic tools, based on the techniques of computer algebra (mainly results on subresultants). In the third section, we study simultaneous inequalities at the real roots of a polynomial and in the fourth section, we consider the coding of real algebraic numbers.

2. Basic Tools and Notations

In the paper, for $P = a_0 X^p + \ldots + a_p$ a polynomial with integer coefficients, we define the norm of P, $N(P)$, by $N(P) = (a_0 + \ldots + a_p^2)^{1/2}$. The size of P is the log of $N(P)$. The length of an integer is the log of the integer. The degree of P is denoted by $\deg(P)$.

Our algorithms are based on a generalization of Sturm theorem, hence on divisions of polynomials, taking care of signs. We therefore require various notations for the signed remainders.

In this section P denotes a polynomial with integer coefficients of degree p and Q denotes a polynomial with integer coefficients of degree q with leading coefficient $b_0 \neq 0$, $q \leq p$.

2.1. SIGNED PSEUDO-REMAINDERS

We denote by $rem(P, Q)$, the remainder of P and Q in the Euclidian division process: so $rem(P, Q)$ has rational coefficients.

0747-7171/90/070039+13 $03.00/0