Using Representation Theorems for Proving Polynomials Non-negative

Salvador Lucas[1]

DSIC, Universidad Politécnica de Valencia
Camino de Vera s/n, 46022 Valencia, Spain

Abstract. Proving polynomials non-negative when variables range on a subset of numbers (e.g., $[0, +\infty)$) is often required in many applications (e.g., in the analysis of *program termination*). Several *representations* for *univariate* polynomials P that are non-negative on $[0, +\infty)$ have been investigated. They can often be used to *characterize* the property, thus providing a method for checking it by trying a *match* of P against the representation. We introduce a new characterization based on viewing polynomials P as *vectors*, and find the appropriate *polynomial basis* $\mathcal{B}$ in which the non-negativeness of the *coordinates* $[P]_\mathcal{B}$ representing P in $\mathcal{B}$ witnesses that P is non-negative on $[0, +\infty)$. Matching a polynomial against a representation provides a way to transform universal sentences $\forall x \in [0, +\infty)\ P(x) \geq 0$ into a *constraint solving* problem which can be solved by using efficient methods. We consider different approaches to solve both kind of problems and provide a quantitative evaluation of performance that points to an early result by Pólya and Szegö's as an appropriate basis for implementations in most cases.

Keywords: Polynomial constraints, positive polynomials, representation theorems.

1 Introduction

Representations of univariate polynomials that are positive ($Pd(I)$) or non-negative ($Psd(I)$) on an *interval* I of real numbers have been investigated (see [14] for a survey) and some of them are useful to *check* the property. In this paper we investigate this question: *which technique is worth to be implemented for a practical use?* Our specific motivation is the development of *efficient* and *automatic* tools for proving *termination* of programs, where polynomials play a prominent role (see [8, 12], for instance) and the focus is on $Psd([0, +\infty))$.

We decompose the whole problem into two main steps: (1) the use of representation theorems to obtain a set of *existential constraints* whose satisfaction witnesses that $(\forall x \geq 0)\ P \geq 0$ holds and (2) the use of constraint solving techniques to obtain appropriate solutions. With regard to (1), several researchers (starting with Hilbert) addressed this problem and contributed in different ways (see Section 2). In this setting, the following test is often used in practice [10]: a polynomial P is $Psd([0, +\infty)^n)$ if *all coefficients of the monomials in P are non-negative*. This has obvious limitations. For instance, $Q(x) = x^3 - 4x^2 + 6x + 1$

is $Psd([0, +\infty))$, but contains negative coefficients. The following observation generalizes this approach (Section 3): $P \in \mathbb{R}[X]$ of degree n can be *represented* as a *vector* $[P]_{\mathcal{B}} = (\alpha_0, \ldots, \alpha_n)^T$ of $n+1$ *coordinates* with respect to a *basis* $\mathcal{B} = \{v_0, \ldots, v_n\} \subseteq \mathbb{R}[X]$, i.e., $P = \alpha_0 v_0 + \alpha_1 v_1 + \cdots + \alpha_n v_n$. Then, P is $Psd([0, +\infty))$ if (i) $[P]_{\mathcal{B}} \geq \mathbf{0}$ and *(ii)* $v_0, \ldots, v_n$ are $Psd([0, +\infty))$. Requiring all coefficients in the representation $P = \sum_{i=0}^{n} p_i x^i$ to be non-negative corresponds to considering the *standard* basis $\mathcal{S}_n = \{1, x, ..., x^n\}$ for polynomials of degree n. In our running example, $[Q]_{\mathcal{S}_3} = (1, 6, -4, 1)^T \not\geq \mathbf{0}$. We define a *parametric* polynomial basis $\mathcal{P}_n$ such that, for all $P \in \mathbb{R}[X]$ of degree n which is $Psd([0, +\infty))$, $[P]_{\mathcal{B}} \geq \mathbf{0}$ for some specific $\mathcal{B}$ which is obtained from $\mathcal{P}_n$ by giving appropriate values to the parameters. We also show how to give value to the parameters.

Example 1. The representation of $Q(x) = x^3 - 4x^2 + 6x + 1$ with respect to $\mathcal{B} = \{1, x, x^2, x(x-2)^2\}$ is $[Q]_{\mathcal{B}} = (1, 2, 0, 1)^T \geq \mathbf{0}$.

Regarding (2), in Section 4 we use a recent, efficient procedure to solve polynomial constraints over finite domains [5] as a reference to provide a quantitative analysis of the characterizations discussed in Sections 2 and 3 and provide an answer to our question. Section 5 discusses some related work and concludes.

2 Representation of polynomials non-negative in $[0, +\infty)$

We consider the following representations of $Psd([0, +\infty))$ polynomials P (see [14]): (1) Hilbert [9]; (2) Pólya and Szegö [13]; (3) Karlin and Studden [11]; and (4) Hilbert's approach using Gram matrices [7].

Remark 1. Our motivation for considering these particular methods is that, in automatic proofs of termination, polynomials P whose non-negativity must be guaranteed are *parametric*, i.e., the coefficients are *not* numbers but rather variables whose value is *generated* by a constraint solving process. All previous methods fit the requirement of being amenable to this practical setting.

We briefly discuss how to use these four methods and also give some cost indicators: $V(n)$ is the number of *parameters* used to match P (of degree n) against the representation, and $I(n)$ is the number of *(in)equalities* which are obtained. The following fact is used later.

Proposition 1. *Let $P, Q \in \mathbb{R}[X_1, \ldots, X_n]$ be $P = \sum_\alpha a_\alpha X^\alpha$ and $Q = \sum_\alpha b_\alpha X^\alpha$. If $a_\alpha \geq b_\alpha$ for all $\alpha \in \mathbb{N}^n$ and Q is $Psd([0, +\infty)^n)$, then P is $Psd([0, +\infty)^n)$.*

In the following, $\div$ and % denote the integer division and remainder, respectively. We say that a polynomial P is a *sum of squares* (or just *sos*, often denoted as $P \in \sum \mathbb{R}[\boldsymbol{X}]^2$) if can be written $P = \sum_i f_i^2$ for polynomials f_i.

2.1 Hilbert

Since $P \in \mathbb{R}[X_1, \ldots, X_n]$ is $Psd([0, +\infty)^n)$ if and only if $H(X_1, \ldots, X_n) = P(X_1^2, \ldots, X_n^2)$ is $Psd(\mathbb{R}^n)$ (note that this transformation *doubles* the degree of P), we can use the following result.

Proposition 2 (Hilbert). [9] *If $P \in \mathbb{R}[X]$ is $Psd(\mathbb{R})$, then P is a sum of two squares of polynomials.*

Example 2. Consider $H(x) = Q(x^2) = x^6 - 4x^4 + 6x^2 + 1 = f_1(x) + f_2(x)$ where $f_i(x) = (a_i x^3 + b_i x^2 + c_i x + d_i)^2$ for $i = 1, 2$. Then, $H(x)$ should match

$\sum_{i=1}^{2} a_i^2 x^6 + 2a_i b_i x^5 + (b_i^2 + 2a_i c_i)x^4 + 2(b_i c_i + a_i d_i)x^3 + (2b_i d_i + c_i^2)x^2 + 2c_i d_i x + d_i^2$

which amounts at *solving* the following *equalities*:

$$\sum_{i=1}^{2} a_i^2 = 1 \qquad \sum_{i=1}^{2} a_i b_i = 0 \qquad \sum_{i=1}^{2} b_i^2 + 2a_i c_i = -4$$
$$\sum_{i=1}^{2} b_i c_i + a_i d_i = 0 \qquad \sum_{i=1}^{2} 2b_i d_i + c_i^2 = 6 \qquad \sum_{i=1}^{2} c_i d_i = 0 \qquad \sum_{i=1}^{2} d_i^2 = 1$$

A solution (with *irrational* numbers) is obtained by using, e.g., *Mathematica*.

We have $V(n) = 2n + 2$ and $I(n) = 2n + 1$.

2.2 Pólya and Szegö

Proposition 3 (Pólya & Szegö). [13] *If P is $Psd([0, +\infty))$, then there are sos polynomials f, g such that $P(x) = f(x) + xg(x)$ and $\deg(f), \deg(xg) \leq \deg(P)$.*

If $f, g \in \sum \mathbb{R}[X]^2$, then both f and xg are $Psd([0, +\infty))$. Thus, Pólya and Szegö's representation actually provides a *characterization*. We can use it, then, to *prove* that P is $Psd([0, +\infty))$ iff P *matches* the representation. Since every univariate *sos* polynomial f can be written as a sum of *two* squares of polynomials, in Proposition 3 we assume $f = f_1^2 + f_2^2$ and $g = g_1^2 + g_2^2$, for polynomials f_i and g_i, $i = 1, 2$. If $n = \deg(P) = 1$, then, since $\deg(f), \deg(xg) \leq 1$, $f, g \in \sum \mathbb{R}[X]^2$ must be *constant* polynomials $f = f_0$ and $g = g_0$. If $n = 2$, then, since $\deg(xg) \leq 2$, $g \in \sum \mathbb{R}[X]^2$ must be a *constant*. If $n > 2$, then $\deg(f_i) = d_1 \leq \lfloor \frac{n}{2} \rfloor$, and $\deg(g_i) = d_2 \leq \lfloor \frac{n-1}{2} \rfloor$. Write $f_i = a_{i,d_1} x^{d_1} + \cdots + a_{i,1} x + a_{i,0}$ and $g_i = b_{i,d_2} x^{d_2} + \cdots + b_{i,1} x + b_{i,0}$ for $i = 1, 2$. Try to *match* the coefficients of the *target* polynomial P against this representation.

Example 3. For our running example Q, we have

$$Q(x) = x^3 - 4x^2 + 6x + 1 = f_1(x) + f_2(x) + x(g_1(x) + g_2(x))$$

where $f_i(x) = (a_i x + b_i)^2$ and $g_i(x) = (c_i x + d_i)^2$ for $i = 1, 2$. Then,

$Q(x) = (c_1^2 + c_2^2)x^3 + (a_1^2 + a_2^2 + 2c_1 d_1 + 2c_2 d_2)x^2 + (2a_1 b_1 + 2a_2 b_2 + d_1^2 + d_2^2)x + b_1^2 + b_2^2$

By Proposition 1, rather than equalities, we solve now the *inequalities*[1]:

$1 \geq c_1^2 + c_2^2;\ -4 \geq a_1^2 + a_2^2 + 2c_1 d_1 + 2c_2 d_2;\ 6 \geq 2a_1 b_1 + 2a_2 b_2 + d_1^2 + d_2^2;\ 1 \geq b_1^2 + b_2^2.$

with: $a_1 = 0$, $a_2 = 0$, $b_1 = 1$, $b_2 = 0$, $c_1 = 1$, $c_2 = 0$, $d_1 = -2$, and $d_2 = 1$.

Each f_i and g_i contributes with $d_1 + 1$ and $d_2 + 1$ parametric coefficients, respectively, i.e., $V(n) = 2(d_1 + 1 + d_2 + 1) = 2(2 + d_1 + d_2) = 2(n + 1) = 2n + 2$. The number of *inequalities* to be solved is $I(n) = n + 1$ (one per coefficient p_i of P).

[1] Using inequalities makes the constraint solving process more flexible and often avoids the use of *irrational* numbers, often out of the scope for most constraint solving tools.

2.3 Karlin and Studden

Theorem 1 (Karlin and Studden). [11, Corollary V.8.1] *Let P_{2m} be a polynomial of degree $2m$ for some $m \geq 0$ with leading coefficient $a_{2m} > 0$. If P_{2m} is $Pd([0, +\infty))$, then there exists a unique representation*

$$P_{2m}(X) = a_{2m} \prod_{j=1}^{m} (X - \alpha_j)^2 + \beta X \prod_{j=2}^{m} (X - \gamma_j)^2$$

where $\beta > 0$ and $0 = \gamma_1 < \alpha_1 < \gamma_2 < \cdots < \gamma_m < \alpha_m < \infty$. Similarly, if P_{2m+1} is a polynomial of degree $2m+1$ for some $m \geq 0$, with leading coefficient $a_{2m+1} > 0$ and P_{2m+1} is $Pd([0, +\infty))$, then there exists a unique representation

$$P_{2m+1}(X) = a_{2m+1} X \prod_{j=2}^{m+1} (X - \alpha_j)^2 + \beta \prod_{j=1}^{m} (X - \gamma_j)^2$$

where $\beta > 0$ and $0 = \alpha_1 < \gamma_1 < \alpha_2 < \gamma_2 < \cdots < \gamma_m < \alpha_{m+1} < \infty$.

Unfortunately, this representation *cannot* be used to *prove* that P is $Pd([0, +\infty))$ by *matching*. For instance, $P = (x-1)^2$ matches it, but it is *not* $Pd([0, +\infty))$. However, Karlin and Studden's representation can be used to prove P to be $Psd([0, +\infty))$ by *matching* if we just require $\alpha_j, \beta, \gamma_j \geq 0$.

Example 4. Since the degree of Q is odd, we let

$$K_Q(x) = x(x - \alpha_2)^2 + \beta(x - \gamma_1)^2 = x^3 + (\beta - 2\alpha_2)x^2 + (\alpha_2^2 - 2\beta\gamma_1)x + \beta\gamma_1^2$$

Thus, we have the following constraints (using Proposition 1):

$$-4 \geq \beta - 2\alpha_2 \quad 1 \geq 0 \quad 6 \geq \alpha_2^2 - 2\beta\gamma_1 \quad 1 \geq \beta\gamma_1^2 \quad \beta \geq 0 \quad \gamma_1 \geq 0 \quad \alpha_2 \geq 0$$

The assignment $\alpha_2 = \frac{9}{4}$, $\beta = \frac{1}{4}$, and $\gamma_1 = \frac{1}{2}$ solves the system.

We have $V(n) = n$ and $I(n) = n + 1 + V(n) = 2n + 1$.

2.4 Hilbert with Gram matrices

An alternative way to use Hilbert's representation is the following.

Theorem 2. [7] *Let P be a polynomial of degree $2m$ and $z(X)$ be the vector of all monomials X^α such that $|\alpha| \leq m$. Then, P is a sum of squares in $\mathbb{R}[X]$ if and only if there exists a real, symmetric, psd matrix B such that $P = z(X)^T B z(X)$.*

Proving $H(x) = P(x^2)$ of degree $2n$ to be *sos* amounts at (1) matching H against $z(X)^T B z(X)$ (where $z(X) = (1, X, \ldots, X^n)^T$) and (2) proving $B \in \mathbb{R}^{n+1 \times n+1}$ positive semidefinite. Since B is symmetric, we need $\frac{(n+1)(n+2)}{2}$ parameters b_{ij} to represent B. Then, we need to solve $2n + 1$ equations in $\frac{(n+1)(n+2)}{2}$ variables (the parameters b_{ij}) corresponding to the monomials in H. According to [15], this can be done by taking $\frac{(n+1)(n+2)}{2} - (2n + 1) = \frac{n^2 - n}{2}$ of the b_{ij} as *unknowns* which can be given appropriate values that are obtained using (2), i.e., B must be positive semidefinite. This can be done by computing the characteristic polynomial $det(zI_{n+1} - B) = \sum_{i=0}^{n} c_i z^i$ of B and requiring its roots to be non-negative [15]. They show that this can be achieved by imposing $(-1)^{i+n+1} c_i \geq 0$ for all $0 \leq i \leq n$. Thus, $V(n) = \frac{(n+1)(n+2)}{2}$. and $I(n) = (2n + 1) + (n + 1) = 3n + 2$.

3 Checking positiveness of polynomials as vectors

Let V be an n-dimensional vector space over the reals and $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis for V. For all n-tuples $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ we write $\boldsymbol{\alpha} \geq 0$ if $\alpha_i \geq 0$ and $\boldsymbol{\alpha} > 0$ if $\alpha_1 > 0$ and $\alpha_2, \ldots, \alpha_n \geq 0$. Every $\boldsymbol{v} \in \mathsf{V}$ can be represented as a *coordinate* vector $[\boldsymbol{v}]_{\mathcal{B}} = (\alpha_1, \ldots, \alpha_n)^T \in \mathbb{R}^n$ such that $\boldsymbol{v} = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Given bases $\mathcal{B}$ and $\mathcal{B}'$ for V, there is a *change of base* matrix (cb-matrix) $M_{\mathcal{B}' \mapsto \mathcal{B}}$ (or just M) which can be used to obtain the *coordinate* representation $[\boldsymbol{v}]_{\mathcal{B}}$ of $\boldsymbol{v}$ in $\mathcal{B}$ from the representation $[\boldsymbol{v}]_{\mathcal{B}'}$ of $\boldsymbol{v}$ in $\mathcal{B}'$: $[\boldsymbol{v}]_{\mathcal{B}} = M[\boldsymbol{v}]_{\mathcal{B}'}$. The set $\mathbb{P}_n$ of *univariate* polynomials of degree at most n is a vectorial space of dimension $n+1$ and has a *standard basis* $\mathcal{S}_n = \{1, x, \ldots, x^n\}$. If $\mathcal{B} = \{v_0, \ldots, v_n\}$ is a basis for $\mathbb{P}_n$ and every $v \in \mathcal{B}$ is $Psd([0, +\infty))$, then given $P \in \mathbb{P}_n$, if $[P]_{\mathcal{B}} = (\alpha_0, \ldots, \alpha_n)^T \geq \mathbf{0}$, then P is $Psd([0, +\infty))$. If $P = \sum_{i=0}^{n} p_i x^i$, this is translated into the *search* of a *basis* $\mathcal{B}$ satisfying the conditions above and a *cb*-matrix $M = M_{\mathcal{S}_n \mapsto \mathcal{B}}$ such that $M\,[P]_{\mathcal{S}_n} \geq \mathbf{0}$. We consider *parametric* bases $\mathcal{B}$ consisting of polynomials with *parametric* coefficients which can be given appropriate values as to fit the requirements above. By a *parametric* polynomial we mean a polynomial $P \in \mathbb{R}[\gamma_1, \ldots, \gamma_k][X]$ over X whose monomials have *coefficients* in $\mathbb{R}[\gamma_1, \ldots, \gamma_k]$; variables $\gamma_1, \ldots, \gamma_k$ are called *parameters*. For all $i \in \mathbb{N}$, consider the parametric univariate polynomials, :

$$\mathsf{P}_i(x) = \prod_{j=1}^{\frac{i}{2}} (x - \gamma_{ij})^2 \text{ if } i \text{ is even} \qquad \mathsf{P}_i(x) = x \prod_{j=1}^{\frac{i-1}{2}} (x - \gamma_{ij})^2 \text{ if } i \text{ is odd}$$

where the empty product is 1, and γ_{ij} are *parameters* satisfying $\gamma_{ij} \geq 0$. For instance, $\mathsf{P}_0(x) = 1$, $\mathsf{P}_1(x) = x$, $\mathsf{P}_2(x) = (x - \gamma_{21})^2 = \gamma_{21}^2 - 2\gamma_{21}x + x^2$, and $\mathsf{P}_3(x) = x(x - \gamma_{31})^2 = \gamma_{31}^2 x - 2\gamma_{31}x^2 + x^3$. Note that for all $i \geq 0$ and $x \geq 0$, $\mathsf{P}_i(x) \geq 0$ and $\mathsf{P}_0(x) > 0$. Given $n \in \mathbb{N}$, let $\mathcal{P}_n = \{\mathsf{P}_0(x), \ldots, \mathsf{P}_n(x)\}$ *ordered* by the sequence $0, 1, \ldots, n$. $\mathcal{P}_n$ is a basis of $\mathbb{P}_n$; this is a consequence of the following.

Theorem 3. *Let $\mathcal{P} = \{P_0, \ldots, P_n\}$ be a set of $n+1$ polynomials such that $P_0 \in \mathbb{R} - \{0\}$ and $deg(P_i) = i$ for all $1 \leq i \leq n$. Then, $\mathcal{P}$ is a basis of $\mathbb{P}_n(x)$.*

Note that $\mathcal{P}_{n+1} = \mathcal{P}_n \cup \mathsf{P}_{n+1}(x)$.

Proposition 4 (Number of parameters in the basis). *Given $n \in \mathbb{N}$, the number $N(n)$ of parameters in $\mathcal{P}_n$ is given by $N(0) = 0$ and $N(n) = N(n-1) + \lfloor \frac{n}{2} \rfloor$ for $n > 0$. Furthermore, $N(n) = \frac{n^2}{4}$ if n is even and $\frac{n^2-1}{4}$ otherwise.*

We prove that $\mathcal{P}_n$ *characterizes* $Psd([0, +\infty))$ and $Pd([0, +\infty))$.

Theorem 4. *A polynomial $P \in \mathbb{R}[X]$ of degree n is $Psd([0, +\infty))$ ($Pd([0, +\infty))$) if and only if $[P]_{\mathcal{P}_n} \geq \mathbf{0}$ (resp. $[P]_{\mathcal{P}_n} > \mathbf{0}$) for some assignment of values $\gamma_{ij} \geq 0$ to the parameters in $\mathcal{P}_n$.*

We show how to compute the cb-matrix $M_n = M_{\mathcal{S}_n \mapsto \mathcal{P}_n}$ for obtaining the representation $[P]_{\mathcal{P}_n} = M_n [P]_{\mathcal{S}_n}$ of $P \in \mathbb{P}_n$ which is required in Theorem 4. In the following, $[\mathsf{P}_n(x)]_{\mathcal{S}_n}^{1,\cdots,n}$ is the n-dimensional vector containing the first n (parametric) coordinates of $[\mathsf{P}_n(x)]_{\mathcal{S}_n}$ (the last one is 1, corresponding to x^n).

Theorem 5 (Incremental cb-matrix). *We have $M_0 = I_1$ and for all $n > 0$,*

$$M_n = \begin{pmatrix} M_{n-1} & -M_{n-1}[\mathsf{P}_n(x)]^{1,\dots,n}_{\mathcal{S}_n} \\ \mathbf{0}_{1\times n} & 1 \end{pmatrix}$$

Example 5. Since $M_1 = I_2$, according to Theorem 5, we have:

$$M_2 = \begin{pmatrix} M_1 & -M_1 \begin{pmatrix} \gamma_{21}^2 \\ -2\gamma_{21} \end{pmatrix} \\ \mathbf{0}_{1\times 2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\gamma_{21}^2 \\ 0 & 1 & 2\gamma_{21} \\ 0 & 0 & 1 \end{pmatrix} \text{ and}$$

$$M_3 = \begin{pmatrix} M_2 & -M_2 \begin{pmatrix} 0 \\ \gamma_{31}^2 \\ -2\gamma_{31} \end{pmatrix} \\ \mathbf{0}_{1\times 3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\gamma_{21}^2 & -2\gamma_{21}^2\gamma_{31} \\ 0 & 1 & 2\gamma_{21} & 4\gamma_{21}\gamma_{31} - \gamma_{31}^2 \\ 0 & 0 & 1 & 2\gamma_{31} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

For our running example $[Q]_{\mathcal{S}_3} = (1, 6, -4, 1)^T$, we *impose* $[Q]_{\mathcal{P}_3} = M_3[Q]_{\mathcal{S}_3} > 0$:

$$\begin{pmatrix} 1 & 0 & -\gamma_{21}^2 & -2\gamma_{21}^2\gamma_{31} \\ 0 & 1 & 2\gamma_{21} & 4\gamma_{21}\gamma_{31} - \gamma_{31}^2 \\ 0 & 0 & 1 & 2\gamma_{31} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 6 \\ -4 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 4\gamma_{21}^2 - 2\gamma_{21}^2\gamma_{31} \\ 6 - 8\gamma_{21} + 4\gamma_{21}\gamma_{31} - \gamma_{31}^2 \\ -4 + 2\gamma_{31} \\ 1 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The corresponding existential constraint:
$\gamma_{21}, \gamma_{31} \geq 0,\ 1+4\gamma_{21}^2-2\gamma_{21}^2\gamma_{31} > 0 \wedge 6-8\gamma_{21}+4\gamma_{21}\gamma_{31}-\gamma_{31}^2 \geq 0 \wedge 2\gamma_{31}-4 \geq 0 \wedge 1 > 0$
is satisfied if $\gamma_{21} = 0$ and $\gamma_{31} = 2$, witnessing Q as $pd([0, +\infty))$ through the coordinate representation $[Q]_{\mathcal{P}_3} = (1, 2, 0, 1)^T$ when $\mathcal{P}_3 = \{1, x, x^2, x(x-2)^2\}$.

Note that $V(n) = N(n) = \frac{n^2 - n\%2}{4}$ and $I(n) = n + 1 + V(n) = n + 1 + \frac{n^2-n\%2}{4}$.

Remark 2. If P is a *parametric polynomial* of degree n, then $[P]_{\mathcal{S}_n}$ is an $n+1$-tuple of parameters which are treated by the constraint solving system which obtains the parameters of the basis $\mathcal{P}_n$ in the same way (see Remark 1).

4 Quantitative analysis

In constraint solving, the number of variables occurring in the whole set of constraints usually dominates the *temporal cost* to reach a solution. In our setting, assuming P of degree n, for each representation method $V(n)$ and $I(n)$ (see Section 2) are as follows:

Method:	Hilbert	P&S	K&S	Gram	*Vector*
$V(n)$**:**	$2n+2$	$2n+2$	$n+1$	$\frac{(n+1)(n+2)}{2}$	$\frac{n^2-n\%2}{4}$
$I(n)$**:**	$2n+1$	$n+1$	$2n+1$	3n+2	$n+1+\frac{n^2-n\%2}{4}$

This table suggests the following conclusion: *for proving Psd*$([0, +\infty))$*, Karlin & Studden is the best choice.* However, this does *not* pay attention to the subsequent *constraint solving* process that we need to use in any implementation. In [5] an efficient procedure to solve polynomial constraints C (e.g., $P \geq 0$, where P is written as a sum of monomials with the corresponding coefficients) is given. The procedure *transforms* a polynomial constraint into a formula of the *linear arithmetic* and then fast, highly efficient *Satisfiability Modulo Theories* (SMT) techniques are used to find a solution. In linear arithmetic (logic) only constants c or additions of linear expressions $c \cdot v$ are allowed, the *atoms* consist of expressions $\ell \bowtie \ell'$ where ℓ, ℓ' are constants or linear expressions and $\bowtie \in \{=, >, \geq\}$, and the *formulas* are combinations of atoms using $\rightarrow$ (implication) and $\wedge$ (conjunction). An initial *preprocessing* $L0$ transforms $P \bowtie 0$ into $\ell_P \bowtie 0$, where ℓ_P is obtained from P by replacing the nonlinear monomials M by new variables x_M; then new atoms $x_M = M$ are added and they are subsequently transformed after further linearization using the following rules, where D is a finite domain of numbers[2]:

Definition 1. *Let C be a pure non-linear constraint and D be a finite set. The transformation rules are the following (where v is a variable):*

L1: $C \wedge x = v^p \Longrightarrow C \wedge \bigwedge_{a \in D}(v = a \rightarrow x = a^p)$, *if* $p > 1$
L2: $C \wedge x = v^p \cdot w \Longrightarrow C \wedge \bigwedge_{a \in D}(v = a \rightarrow x = a^p \cdot w)$
L3: $C \wedge x = v^p \cdot M \Longrightarrow C \wedge \bigwedge_{a \in D}(v = a \rightarrow x = a^p \cdot x_M) \wedge\ x_M = M$
if M is not linear and v does not occur in M

For $x = M_0$ where M_0 is a monomial with m different variables, if M_0 consists of at most two variables, one of them of degree 1, then $L1$ or $L2$ apply; no new variables are introduced and the equality is transformed into $|D|$ new linear formulas. If $M_0 = v^p M$ contains m variables and M is not linear, then only $L3$ applies, and then introduces a *new* variable x_M together with $|D|$ new linear formulas and a new equality $x_M = M$ where M has $m - 1$ variables.

Example 6. For instance, for $1 \geq c_1^2 + c_2^2$ in Example 3,

$$\begin{aligned} 1 \geq c_1^2 + c_2^2 &\rightsquigarrow_{L0} 1 \geq x_{c_1^2} + x_{c_2^2} \wedge x_{c_1^2} = c_1^2 \wedge x_{c_2^2} = c_2^2 \\ &\rightsquigarrow_{L1} 1 \geq x_{c_1^2} + x_{c_2^2} \wedge \textstyle\bigwedge_{d \in D} c_1 = d \rightarrow x_{c_1^2} = d^2 \wedge \bigwedge_{d \in D} c_2 = d \rightarrow x_{c_2^2} = d^2 \end{aligned}$$

we obtain $1 + 2|D|$ linear formulas and 2 new variables are required.

In the following, $V_L(n)$ is the number of new variables introduced by $L0$. And if P is the targeted polynomial, p_i for $0 \leq i \leq n$ is the coefficient of x^i in P.

Hilbert. If $f = \sum_{j=0}^{d} f_j x^j$ is a parametric polynomial of degree $d > 0$, then the coefficient c_i of x^i in f^2 is obtained from the products $f_r f_s$ such that $r + s = i$. Here, $f_s f_r$ does *not* count as a new combination because $f_r f_s + f_s f_r = 2 f_r f_s$. If $i \leq d$ we have different contributing combinations from $(0, i)$ to $(i \div 2, i - i \div 2)$, i.e., $1 + i \div 2$ combinations. If $i > d$, then we have different contributing combinations starting from $(d - i, d)$, i.e., $1 + (2d - i) \div 2 = 1 + d - i \div 2 - i\%2$ combinations. Overall, if $\mu_d(c_i) = 1 + i \div 2$, if $i \leq d$, and $\mu_d(c_i) = 1 + d - i \div 2 - i\%2$, if $i > d$, then c_i consists of a sum of $\mu_d(c_i)$ monomials $f_r f_s$ all of them of degree 2.

[2] Simplified definition which only uses a *single* domain of values for all variables.

When matching $P(x) = \sum_{i=0}^{n} p_i x_i^i$ against Hilbert's representation, each p_i, $0 \leq i \leq n$ is matched by a sum c_{2i} of $2\mu_n(c_{2i})$ expressions of degree 2 (in the parameters). However, for all $0 \leq i < n$, there are *additional* equations $c_{2i+1} = 0$ which are due to the *duplication* of the degree of P before the matching. Therefore, there are $2n+1$ equations gathering

$$\sum_{i=0}^{n} 2\mu_n(c_{2i}) + \sum_{i=0}^{n-1} 2\mu_n(c_{2i+1}) = 2\left(\sum_{i=0}^{n} \mu_n(c_{2i}) + \sum_{i=0}^{n-1} \mu_n(c_{2i+1})\right)$$

quadratic terms all together, i.e., $V_L(n) = 2\left(\sum_{i=0}^{n} \mu_n(c_{2i}) + \sum_{i=0}^{n-1} \mu_n(c_{2i+1})\right)$.

Polya & Szegö. When matching $P = \sum_{i=0}^{n} p_i x_i^i$ against Polya and Szegö's representation in Section 2.2, if $n = 1$, then p_0 and p_1 are matched to *squared* constants f_0^2 and g_0^2, respectively. If $n = 2$, then p_1 is matched to a sum of *two* monomials of degree 2 each; finally, if $n \geq 3$, then p_0 and p_n are each of them matched to a sum of 2 squares, and each p_i, $0 < i < n$ is matched to a sum of $2\mu_{n\div 2}(c_i) + 2\mu_{(n-1)\div 2}(c_{i-1})$ expressions which are parametric coefficients: the coefficients of monomials of degree i from f_1^2 and f_2^2, and the coefficients of monomials of degree $i-1$ from g_1^2 and g_2^2. All these parametric coefficients have degree 2. We have *two* equations with *two* terms and $n-1$ equations gathering

$$\begin{aligned}\sum_{i=1}^{n-1} 2\mu_{n\div 2}(c_i) + 2\mu_{(n-1)\div 2}(c_{i-1}) &= 2\left(\sum_{i=1}^{n-1} \mu_{n\div 2}(c_i) + \sum_{i=1}^{n-1} \mu_{(n-1)\div 2}(c_{i-1})\right) \\ &= 2\left(1 + \mu_{n\div 2}(c_{n-1}) + \sum_{i=1}^{n-2} \mu_{n\div 2}(c_i) + \mu_{(n-1)\div 2}(c_i)\right)\end{aligned}$$

terms. Terms M of degree 2 require a new variable x_M in the initial step $L0$. Overall, $V_L(1) = 2$, $V_L(2) = 3 \cdot 2 = 6$ and, for $n \geq 3$:

$$V_L(n) = 6 + 2\left(\mu_{n\div 2}(c_{n-1}) + \sum_{i=1}^{n-2} \mu_{n\div 2}(c_i) + \mu_{(n-1)\div 2}(c_i)\right)$$

Karlin & Studden. If $\alpha \in \{0, \ldots, n\}^m$, we let $|\alpha| = \sum_{i=1}^{m} \alpha_i$. Note that $(\prod_{i=1}^{m}(x - a_i))^n = \sum_{i=0}^{mn}(-1)^i(\sum_{\alpha\in\{0,\ldots,n\}^m, |\alpha|=mn-i} \boldsymbol{a}^\alpha)x^i$. If $n = 1$, there are $\binom{m}{m-i} = \binom{m}{i}$ *parametric* monomials $\boldsymbol{a}^\alpha$ (all of them of degree $m-i$ with respect to parameters a_i) accompanying x^i. If $n = 2$, we can obtain the number of monomials accompanying x^i as follows. There are $\binom{m}{p}$ monomials $\boldsymbol{a}^\alpha$ with $\alpha \in \{0,1\}^m$ and $|\alpha| = m - p$. Here, $0 \leq p \leq m$. These monomials can contribute to a monomial of degree $2m-i$ for x^i. However, note that only those monomials satisfying $m-p \leq 2m-i$ (i.e., $p \geq i-m$) will be useful; otherwise, the monomials $\boldsymbol{a}^\alpha$ *exceed* the required degree $2m-i$ for x^i. If we replace $2m - i - (m - p) = m - i + p$ occurrences of 1 by 2 in α to yield α' (with $m-p-(m-i+p) = i-2p$ occurrences of 1 only), then, $|\alpha'| = 2(m-i+p)+i-2p = 2m-i$ as desired. We can do that in $\binom{m-p}{m-i+p}$ different ways. However, this process makes sense only if α has *enough* occurrences of 1, i.e., if $2(m-p) \geq 2m-i$ (equivalently, $2p \leq i$, i.e., $p \leq i \div 2$) so that the replacement of occurrences of 1 by 2 in α actually leads to the appropriate α'. Overall, x^i comes with a parametric coefficient of

$$\mathsf{mon}(m,i) = \sum_{p=max(0,i-m)}^{i\div 2} \binom{m}{p}\binom{m-p}{m-i+p}$$

monomials of degree $2m-i$ (in the parameters a_i).

When matching a polynomial P of degree $2m$ against Karlin & Studden representation, we get $2m+1$ constraints $C_i \leq p_i$, $0 \leq i \leq 2m$, where C_i consists of $\mathsf{mon}(m,i)$ monomials of degree $2m-i$ (coming from the first term of $P_{2m}(X)$ in Theorem 1) and $\mathsf{mon}(m-1,i-1)$ monomials of degree $2m-i$ (due to the product with β and X) coming from the second term of $P_{2m}(X)$. Therefore, C_i consists of nonlinear monomials if $2m-i > 1$ (i.e., $i < 2m-1$). Overall, we have $\sum_{i=0}^{2m-2}(\mathsf{mon}(m,i)+\mathsf{mon}(m-1,i-1))$ nonlinear monomials. Similarly, P of degree $2m+1$ yields $2m+2$ constraints $C_i = p_i$, $0 \leq i \leq 2m+1$, where C_i consists of $\mathsf{mon}(m,i-1)$ monomials of degree $2m-i+1$ (coming from the first term of $P_{2m}(X)$ above) and $\mathsf{mon}(m,i)$ monomials of degree $2m-i+1$ (due to the product with β) coming from the second term of $P_{2m}(X)$. Therefore, C_i consists of nonlinear monomials if $2m-i+1 > 1$ (i.e., $i < 2m$). Overall, $\sum_{i=0}^{2m-1}(\mathsf{mon}(m,i-1)+\mathsf{mon}(m,i))$ nonlinear monomials. Hence,

$$V_L(n) = \begin{cases} \sum_{i=0}^{2m-2}(\mathsf{mon}(m,i)+\mathsf{mon}(m-1,i-1)) & \text{if } n=2m \\ \sum_{i=0}^{2m-1}(\mathsf{mon}(m,i-1)+\mathsf{mon}(m,i)) & \text{if } n=2m+1 \end{cases}$$

Vector. In the following, $\mu(e)$ is the *number of monomials* in a *parametric* polynomial expression e in normal form; $\kappa(e)$ is the number of *constant* monomials in e ($\kappa(e) \in \{0,1\}$); $\lambda(e)$ is the number of *linear and non constant monomials* in e ($\lambda(e) \in \{0,1\}$); and $\overline{\lambda}(e)$ is the number of *nonlinear* monomials in e. Clearly, $\mu(e) = \kappa(e) + \lambda(e) + \overline{\lambda}(e)$. Note that, since κ, λ, and $\overline{\lambda}$ are mutually exclusive, identifying $\mu(e)$ with one of them implies that the other are null. Finally, $\delta(e)$ is the *common* degree of all monomials in e (or $\perp$ if it does not exist). A polynomial $\mathsf{P}_n(x)$ consists of parametric coefficients $\pi_{n,i}$ for $0 \leq i \leq n$, where $\pi_{n,n} = 1$ (i.e., $\mu(\pi_{n,n}) = \kappa(\pi_{n,n}) = 1$ and $\delta(\pi_{n,n}) = 0$). If $n > 0$ is *even* ($n = 0$ is a particular case of the previous one), then for all $0 \leq i < n$, $\pi_{n,i}$ consists of a sum of $\mu(\pi_{n,i}) = \mathsf{mon}(n \div 2, i)$ monomials, all of them of degree $n-i$ (i.e., $\delta(\pi_{n,i}) = n-i$). Thus, $\pi_{n,i}$ is linear (and nonconstant) if $n-i = 1$. Therefore, $\mu(\pi_{n,n-1}) = \lambda(\pi_{n,n-1})$ and, for all $0 \leq i < n-1$, $\mu(\pi_{n,i}) = \overline{\lambda}(\pi_{n,i})$ and $\delta(\pi_{n,i}) = n-i$. If n is *odd*, then $\pi_{n,0} = 0$ and for all $0 < i < n$, $\pi_{n,i}$ consists of a sum of $\mu(\pi_{n,i}) = \mathsf{mon}(n \div 2, i-1)$ monomials, all of them of degree $n-i+1$ (i.e., $\delta(\pi_{n,i}) = n-i+1$). Summarizing: $\mu(\pi_{n,i}) = \mathsf{mon}(n \div 2, i-(n\%2))$. A constraint $P \geq 0$ is translated into a set of $n+1$ inequalities $C_i \geq 0$, where C_i is the result of multiplying the i-th row of $M_n = (m^n_{ij})_{n+1\times n+1}$ and $[P]_{\mathcal{S}_n}$, the vector of coefficients of P, for $i = 0,\ldots,n$. We have the following results.

Proposition 5. *For all n, $\mu(m^n_{1,2}) = 0$ and for all $1 \leq j < i \leq n$, $\mu(m^n_{i,i}) = 1$ and $\mu(m^n_{i,j}) = 0$. Let $n > 1$. For all $1 \leq i \leq n$,*

1. $\mu(m^n_{i,n+1}) = \sum_{j=1}^n \mu(m^{n-1}_{ij})\mu(\pi_{n,j-1}) = \sum_{j=1}^n \mu(m^{n-1}_{ij})\mathsf{mon}(n \div 2, (j-1) - n\%2)$.
2. $\delta(m^n_{i,n+1}) = \delta(m^{n-1}_{i,n}) + 1 = n + 1 - i$.

Proposition 6. *$V_L(0) = V_L(1) = 0$ and for all $n > 1$, $V_L(n) = V_L(n-1) + \sum_{i=1}^{n-1} \mu(m^n_{i,n+1})$.*

4.1 Comparison

Let $V_P(n) = V(n) + V_L(n)$ be the number of parameters obtained *after matching a given representation and issuing the preprocessing step L0 for the linearization.* The following table shows $V_P(n)$ for some degrees n of the targeted polynomial P for the considered representation methods[3].

Method	1	2	3	4	5	6	7	8	9	10	20	100
Hilbert	10	18	28	40	54	70	88	108	130	154	504	10504
P&S	6	10	20	28	36	46	56	68	80	94	284	5404
K&S	2	4	7	13	20	38	57	111	166	328	78741	$9.57 \cdot 10^{23}$
Vector	0	2	6	28	96	498	2322	15308	93696	758086	$2.48 \cdot 10^{16}$	$< \infty$

Although the range of values for n is small, the trend for the different methods is clear and suggests that, for $n > 6$, Pólya & Szegö's representation provides the best starting point for an implementation. Let's reason that this is actually the case. Let $W_L(n)$ be the number of variables introduced by the *linearization* after using $L0$ and $L1, \ldots, L3$. Obviously, $V_L(n) \leq W_L(n)$. Let $V_T(n) = V(n) + W_L(n)$ be the number of variables occurring in the linear formula obtained by the linearization process. The number $F_L(n)$ of new formulas introduced by the linearization is bounded by $|D|W_L(n) \leq F_L(n)$. And the total number of formulas is $F_T(n) = I(n) + F_L(n)$, thus bounded by $I(n) + |D|W_L(n) \leq F_T(n)$.

Since the degree of *all monomials* in the parametric polynomials in the representation is 2, for Pólya and Szegö's representation $W_L^{PS}(n) = V_L^{PS}(n)$ (the linearization process will *not* introduce more variables after $L0$). Thus, $V_T^{PS}(n) = V^{PS}(n) + V_L^{PS}(n) = V_P^{PS}(n)$. The $V_L^{PS}(n)$ equations $x_M = M$ are transformed by the *application* of $L1$ or $L2$ only (because $\deg(M) = 2$) into $F_L^{PS}(n) = |D|V_L^{PS}(n)$ new linear formulas. Thus, $F_T^{PS}(n) = I^{PS}(n) + |D|V_L^{PS}(n)$.

Since for $M \in \{Hilbert, KS, Vector, G\}$, $V_T^{PS}(n) = V_P^{PS}(n) < V_P^M \leq V_T^M(n)$ for all $n > 6$ (see the table above[4]), and, since $I^{PS}(n) < I^M(n)$ for all $n > 1$, we have $F_T^{PS}(n) = I^{PS}(n) + |D|V_L^{PS}(n) < I^M(n) + |D|W_L^M(n) \leq F_T^M(n)$ for all $n > 6$, we finally conclude that *Pólya and Szegö's representation is the best choice for an implementation* using the constraint solving method in [5]: it *minimizes* both the number of variables $V_T(n)$ and formulas $F_T(n)$ to be considered.

[3] Obtained using Haskell encodings of the cost formulas in Appendix B.

[4] Although we do not provide information about $V_L^G(n)$, note that $V^G(n)$ and V_T^{PS} are already very similar. Thus, assuming $V_T^{PS}(n) < V_T^G(n)$ is natural.

5 Related work and conclusions

In Section 3, we have shown that the notions of *polynomial bases* and *vector coordinates* can be used instead of that of *monomials* and *monomial coefficients* when testing univariate polynomials P for $Psd([0,+\infty))$ and $Pd([0,+\infty))$. The quantitative analysis in the previous section, though, suggests that this new method is hardly useful in practice. We show its theoretical interest as improving on the use of *Bernstein's polynomials* [3], which inspired our developments.

$Psd([0,+\infty))$ and $Psd([-1,1])$ are related through *Goursat transform* (see [14]): Given $P \in \mathbb{R}[X]$ of degree n, we let $\widetilde{P}(X) = (1+X)^n P(\frac{1-X}{1+X})$. Furthermore, $\widetilde{\widetilde{P}}(X) = 2^n P(X)$. Then, P is $Psd([-1,1])$ if and only if $\widetilde{P}$ is $Psd([0,+\infty))$ and $\deg(\widetilde{P}) \leq n$, see [14, Lemma 1]. Testing $Pd([-1,1])$ or $Psd([-1,1])$ of univariate polynomials $P \in \mathbb{R}[X]$ on $[-1,1]$ can be done by using the so-called Bernstein's basis [6]: if $[P]_{\mathcal{B}_n} > \mathbf{0}$, for the Bernstein basis $\mathcal{B}_n$ (which consists of polynomials of degree n only) then P is $Pd([-1,1])$ [2]. Unfortunately, $\mathcal{B}_n$ does *not* capture all $P \in Pd([-1,1])$ as positive vectors $[P]_{\mathcal{B}_n}$. For instance, $P(X) = 5X^2 - 4X + 1$ is positive on $[-1,1]$ but $[P]_{\mathcal{B}_2} \not> \mathbf{0}$ [6]. Nevertheless, for each $P \in Pd([-1,1])$ of degree n the so-called *Bernstein's Theorem* [4] ensures the existence of some $p \geq n$ such that $[P]_{\mathcal{B}_p}$ consists of *positive coordinates* only (the minimum of those p is called the *Bernstein degree* of P). Unfortunately, such p can be much higher than n. For instance, for $P(X) = 5X^2 - 4X + 1)$ we need to consider 23 polynomials in Bernstein's basis. Even worst, the Bernstein degree of a polynomial P is not usually known, and we have to (over)estimate it. For instance, a the recent estimation [6] is $\frac{n(n-1)}{2}\frac{M}{\lambda}$, where n is the degree of the polynomial, M is the maximum value of the coordinates $[P]_{\mathcal{B}_n}$ of P in the Bernstein basis of degree n, and λ is the minimum of P on $[-1,1]$. For $P(X) = 5X^2 - 4X + 1$ we have $n = 2$, $M = 10$, $\lambda = \frac{1}{5}$, and a estimation of 50, far beyond 23, the real Bernstein degree of P. In [6], this problem is addressed by using *partitions* of $[-1,1]$ where we are able to represent P in a Bernstein basis of degree n by using positive coordinates only. However, we need to produce several (up to $n+1$) partitions of $[-1,1]$, compute the corresponding representations of P, etc. Furthermore, it is unclear how [6] would be used with parametric polynomials (see Remark 1).

Example 7. For our running example, we get $\widetilde{Q}(X) = -10X^3 + 4X^2 + 10X + 4$. According to [6, page 640], for $\mathcal{B}_3 = \{\binom{3}{i}\frac{(1-X)^{3-i}(X+1)^i}{8} \mid 0 \leq i \leq 3\}$, i.e.,

$$\{\frac{1}{8}(1-3x+3x^2-x^3), \frac{3}{8}(1-x-x^2+x^3), \frac{3}{8}(1+x-x^2-x^3), \frac{1}{8}(1+3x+3x^2+x^3)\}$$

we have: $S_{\mathcal{S}_3 \mapsto \mathcal{B}_3} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & -\frac{1}{3} & -\frac{1}{3} & 1 \\ 1 & \frac{1}{3} & -\frac{1}{3} & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ and $[\widetilde{Q}]_{\mathcal{B}_3} = S_{\mathcal{S}_3 \mapsto \mathcal{B}_3}[\widetilde{Q}]_{\mathcal{S}_3} = \begin{pmatrix} 8 \\ -\frac{32}{3} \\ 16 \\ 8 \end{pmatrix}$,
which does *not* witness $\widetilde{Q}$ as $Psd([-1,1])$ due to the negative coordinate $-\frac{32}{3}$ in

$[\widetilde{Q}]_{\mathcal{B}_3}$. The estimated Bernstein degree (for $n = 3$, $M = 16$ and $\lambda \simeq 1.22$) is 40, i.e, a 40-square cb-matrix is required! This can be compared with Example 5.

We have investigated methods for proving univariate polynomials $Psd([0, +\infty))$, and a quantitative evaluation of the requirements needed to make a practical use of them suggests that an early result by Pólya and Szegö's provides an appropriate basis for implementations in most cases. An important motivation and contribution of this work in connection with the development of tools for automatically proving termination is that we avoid the need of explicitly requiring that parametric polynomials arising in proofs of termination have non-negative coefficients (which is the usual practice in termination provers, see [8, 12]). We will use our new findings in future versions of the tool MU-TERM [1].

Acknowledgements. I thank the anonymous referees for their valuable comments.

References

1. B. Alarcón, R. Gutiérrez, S. Lucas, R. Navarro-Marset. Proving Termination Properties with MU-TERM. In *Proc. of AMAST'10*, LNCS 6486:201-208, 2011.
2. S. Basu, R. Pollack, and M.-F. Roy. Algorithms in Real Algebraic Geometry. Springer-Verlag, Berlin, 2006.
3. S. Bernstein. Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités. *Communic. Soc. Math. de Kharkow* 13(2):1-2, 1912.
4. S. Bernstein. Sur la répresentation des polynômes positifs. *Communic. Soc. Math. de Kharkow* 14(2):227-228, 1915.
5. C. Borralleras, S. Lucas, A. Oliveras, E. Rodríguez, and A. Rubio. SAT Modulo Linear Arithmetic for Solving Polynomial Constraints. *Journal of Automated Reasoning*48:107-131, 2012.
6. F. Boudaoud, F. Caruso, and M.-F. Roy. Certificates of Positivity in the Bernstein Basis. *Discrete Computational Geometry* 39:639-655, 2008.
7. M.D. Choi, T.Y. Lam, and B. Reznick. Sums of squares of real polynomials. In *Proc. of the Symposium on Pure Mathematics*, vol. 4, American Mathematical Society, pages 103-126, 1995.
8. E. Contejean, C. Marché, A.-P. Tomás, and X. Urbain. Mechanically proving termination using polynomial interpretations. Journal of Automated Reasoning, 32(4):315-355, 2006.
9. D. Hilbert. Über die Darstellung definiter Formen als Summe von Formenquadraten. Mathematische Annalen 32:342-350, 1888.
10. H. Hong and D. Jakuš. Testing Positiveness of Polynomials. *Journal of Automated Reasoning* 21:23-38, 1998.
11. S. Karlin and W.J. Studden. Tchebycheff systems: with applications in analysis and statistics. Interscience, New York, 1966.
12. S. Lucas. Polynomials over the reals in proofs of termination: from theory to practice. *RAIRO Theoretical Informatics and Applications*, 39(3):547-586, 2005.
13. G. Polya and G. Szegö. Problems and Theorems in Analysis II Springer-Verlag, 1976.
14. V. Powers and B. Reznick. Polynomials that are positive on an interval. *Transactions of the AMS* 352(10):4677-4692, 2000.
15. V. Powers and T. Wörmann. An algorithm for sums of squares of real polynomials. *Journal of Pure and Applied Algebra* 127:99-104, 1998.